2018

LES of flow over bumps and machine learning augmented turbulence modeling

Racheet Matai
Iowa State University

LES of flow over bumps and machine learning augmented turbulence modeling

by

Racheet Matai

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Aerospace Engineering

Program of Study Committee:
Paul Durbin , Major Professor
Alberto Passalacqua
Alric Rothmayer
Jin Tian
Leifur Leifsson

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

DEDICATION

I would like to dedicate this thesis to my parents. Their unwavering love, support and encouragement has helped me endure and enjoy life.

TABLE OF CONTENTS

Page

LIST OF TABLES v

LIST OF FIGURES vi

ACKNOWLEDGMENTS xi

ABSTRACT xii

CHAPTER 1. INTRODUCTION 1

1.1 Simulation of turbulent flows 2

1.1.1 Direct Numerical Simulations 2

1.1.2 Large Eddy Simulations 2

1.1.3 Reynolds Averaged Naiver Stokes Simulations 3

1.2 Overview 4

CHAPTER 2. RANS MODELS AND METHODOLOGY 5

2.1 Two equation closure models 6

2.1.1 k-ϵ model 6

2.1.2 k-ω model 7

2.2 Methodology 8

CHAPTER 3. LES AND FIELD INVERSION 9

3.1 Introduction 10

3.2 Simulation overview 13

3.2.1 Verification 16

3.2.2 Validation 19

3.3 Flow Analysis 24

3.3.1 Base case . . . 24
3.3.2 Bump series . . . 30
3.4 Data Extraction by Inversion . . . 44
3.5 Conclusion . . . 53
CHAPTER 4. ZONAL EDDY VISCOSITY MODEL USING MACHINE LEARNING . . . 56
4.1 Introduction . . . 56
4.2 Methodology . . . 58
4.2.1 Sensitivity Analysis . . . 62
4.2.2 Feature Set Formulation . . . 68
4.3 Results . . . 68
4.3.1 Training Cases . . . 68
4.3.2 Test Cases . . . 71
4.4 Conclusion . . . 77
CHAPTER 5. SUMMARY . . . 79
APPENDIX A. ANN FOR TURBULENCE MODELING . . . 80
A.1 Test Cases . . . 82
A.1.1 h26 . . . 82
A.1.2 Curved back step . . . 84
A.1.3 Limitations . . . 84
APPENDIX B. CONTINUOUS ADJOINT FORMULATION . . . 85

LIST OF TABLES

Page

Table 3.1 Grid characteristics of the bump simulation at $Re_\theta = 2500$ 15

Table 3.2 Grid characteristics of flat plate boundary layer simulation to validate budgets. 18

Table 3.3 Summary of bumps. 30

Table 4.1 Bins for classification of zones 60

Table 4.2 hyper parameters for bag of trees 62

Table 4.3 Bins for classification of zones for modified model 1 64

Table 4.4 Bins for classification of zones for modified model 2 64

Table 4.5 Input Feature set. Note that in OpenFOAM incompressible solvers, p/ρ is computed instead of pressure and hence p here has the units $length^2/Time^2$ 67

Table 4.6 Grid characteristics 69

Table A.1 Input Feature set for ANN. 80

LIST OF FIGURES

Page

Figure 3.1 Mesh used for the LES simulation, shown for the h20 reference bump. Note the disconnected recycling domain for generating the inlet flow (C = 0.305m). 14

Figure 3.2 Grid Convergence for highest bump (h42), LES: —— Mesh 1, - - - Mesh 2, ····· Mesh 3, — - — Mesh 4. RANS:—— $y_+ = 0.5$, - - - $y_+ = 1$, ····· $y_+ = 2$. Note y_+ is at the inlet. 15

Figure 3.3 Ratio of subgrid viscosity to kinematic viscosity for highest bump (h42). 15

Figure 3.4 Ratio of maximum grid spacing to Kolmogorov length scale for highest bump (h42). 16

Figure 3.5 Spanwise correlation function for —— $\overline{uu}$, - - - $\overline{vv}$ and ∘ $\overline{ww}$ at $x/C = -0.26$. 16

Figure 3.6 Comparison of budget of TKE with DNS. T: Turbulent transport, Pr: Pressure transport, D: Resolved Dissipation, A: Advection, P: Production, V: Viscous transport, s: subscript for Spalart's DNS data. 18

Figure 3.7 Time and span averaged x-velocity at various x locations. —— LES, ∗ experiment. 20

Figure 3.8 Time and span averaged $\overline{uu}$ at various x locations. —— LES, —•— single wire, ∗ x-wire. 21

Figure 3.9 Time and span averaged $\overline{vv}$, $\overline{ww}$ and $-\overline{uv}$ at various x locations normalized by U^2_{ref}. —— LES, ∗ experiment. 23

Figure 3.10 Pressure Coefficient, Pressure gradient parameters and intermittent flow reversal at wall. 25

Figure 3.11 Skin Friction Coefficient. 25

Figure 3.12 TKE contours multiplied by $10^2/U_{ref}^2$. 26

Figure 3.13 Production tensor component contours multiplied by $10^2\theta/U_{ref}^3$. Top to bottom: P_{11}, P_{12}, P_{22}, $P_{11}+P_{22}$. C_f vs. x/C is shown for comparison. . 27

Figure 3.14 Velocity gradient components. multiplied by θ/U_{ref}. Top to bottom: $\partial_1 U_1$, $\partial_1 U_2$, $\partial_2 U_1$, $\partial_2 U_2$. C_f vs. x/C is shown for comparison. 28

Figure 3.15 TKE budget at various x locations. Terms multiplied by $10^2\theta/U_{ref}^3$. . 29

Figure 3.16 Bump geometries. 30

Figure 3.17 Comparison of skin friction coefficient and pressure for different bump crest heights. 31

Figure 3.18 Pressure gradient parameters and intermittent flow reversal at wall for bumps with different crest heights. 32

Figure 3.19 Comparison of skin friction coefficient and pressure for different inlet momentum thicknesses (h20). 32

Figure 3.20 Instantaneous Skin Friction Coefficient for h42. 33

Figure 3.21 Time and span averaged x-velocity at various x locations. —— h20, - - - h26, ····· 31, — - — 38, ▬▬ 42. 34

Figure 3.22 Time and span averaged $\overline{uu}$ at various x locations. —— h20, - - - h26, ····· 31, — - — 38, ▬▬ 42. 35

Figure 3.23 Time and span averaged $\overline{vv}$ at various x locations. —— h20, - - - h26, ····· 31, — - — 38, ▬▬ 42. 36

Figure 3.24 Time and span averaged $\overline{ww}$ at various x locations. —— h20, - - - h26, ····· 31, — - — 38, ▬▬ 42. 37

Figure 3.25 Time and span averaged -$\overline{uv}$ at various x locations. —— h20, - - - h26, ····· 31, — - — 38, ▬▬ 42. 38

Figure 3.26 $10^2k/U^2$ at various x locations. —— h20, - - - h26, ····· 31, — · — 38, ▬▬ 42. 39

Figure 3.27 TKE contours multiplied by $10^2/U_{ref}^2$. 40

Figure 3.28 Production tensor component contours for h42 multiplied by $10^2\theta/U_{ref}^3$. Top to bottom: P_{11}, P_{22}, P_{12}, $P_{11}+P_{22}$. C_f vs. x/C is shown for comparison. 41

Figure 3.29 TKE budget for h42 at various x locations. Terms multiplied by $10^2\theta/U_{ref}^3$. 42

Figure 3.30 C_f of inverse solution compared to LES and RANS solutions and the objective function for the optimization (h20). 46

Figure 3.31 Comparison of streamwise velocity at various x locations (h20). $*$ base model, $\circ$ optimized, —— LES. 46

Figure 3.32 correction profiles (β) at various x locations (h20). 47

Figure 3.33 Comparison of ν_t at various x locations (h20). $*$ base model, $\circ$ optimized. 47

Figure 3.34 Comparison of k at various x locations (h20). 48

Figure 3.35 Comparison of inverse solutions with different base models and with the correction factor multiplied to the destruction of ω team instead of production. (h20). $*$: base k-ω model, $\diamond$: base k-ω SST, $\circ$: β multiplied to production of ω of base k-ω model, $+$: β multiplied to destruction of ω of base k-ω model, $\times$: β multiplied to production of ω of base k-ω SST model. 48

Figure 3.36 Comparison of streamwise velocity at various x locations for inverse solutions with different cost functions (h20). 49

Figure 3.37 Comparison of ν_t at various x locations for inverse solutions with different cost functions (h20). 50

Figure 3.38 C_f of inverse solution compared to LES and RANS solutions and the objective function for the optimization (h38). 51

Figure 3.39 Comparison of streamwise velocity at various x locations (h38). . . . 52

Figure 3.40 Comparison of ν_t at various x locations (h38). $*$ base model, $\circ$ optimized, —— least squares. 53

Figure 3.41 Comparison of k for RANS and inverse solutions; terms multiplied by $10^2/U_{ref}^2$. 54

Figure 4.1 Comparison of skin friction coefficient and pressure for different bump crest heights 61

Figure 4.2 Comparison of skin friction coefficient and pressure for different bump crest heights 62

Figure 4.3 Sensitivity of C_f to bin width on training (h20) and test (h42) data. 63

Figure 4.4 Sensitivity of C_f to number of learners and number of splits (h42). . 66

Figure 4.5 Models trained on same data with same hyper parameters (h42). . . 66

Figure 4.6 Estimated importance of features. 67

Figure 4.7 Skin Friction Coefficient and zones for h20. 69

Figure 4.8 x-velocity at various x locations for h20. 70

Figure 4.9 Skin Friction Coefficient and zones for h38. 70

Figure 4.10 x-velocity at various x locations for h38. 71

Figure 4.11 velocity profile for channel flow at $Re_\tau = 5,500$ and flat plate at $Re_\theta = 1,200$. 72

Figure 4.12 Back Step C_f and zones. 73

Figure 4.13 Back Step experimental data compared with base model and ML model.. 73

Figure 4.14 Obi diffuser experimental data compared with base model and ML model. 74

Figure 4.15 h42 skin friction and zones. 75

Figure 4.16 x-velocity at various x locations for h42. 76

Figure 4.17 Skin Friction Coefficient and zones for NASA Hump. 77

Figure A.1 Skin friction coefficient for h20. 81
Figure A.2 x-velocity at various x locations for h20. 82
Figure A.3 Skin friction coefficient for h26. 83
Figure A.4 x-velocity at various x locations for h26. 83
Figure A.5 Skin friction coefficient for curved back step. 84

Figure B.1 Skin friction coefficient for h20. 87
Figure B.2 x-velocity at various x locations for h20. 87

ACKNOWLEDGMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis. First and foremost, Dr. Paul Durbin for his guidance, patience and support throughout this research and the writing of this thesis. His insights and words of encouragement have often inspired me. I would also like to thank my committee members for their contributions to this work and Dr. Karthik Duraisamy for providing us with the optimization code.

Moreover, I would like to thank Dr. Saleh Nabi and Dr. Piyush Grover who played a crucial role in helping me implement the continuous adjoint solver.

Additionally, I would like to thank my friends and labmates Rajarshi, Joel, Umair, Rikhi, Zifei, Vishal and Shujaut for not only sharing their knowledge, but also adding much needed levity to the everyday lab environment.

Finally, I would like to thank Jagan, James and Evan amongst many others in the skydiving community for helping me experience true *"human flight"*.

ABSTRACT

Predicting drag over complex bodies plays a crucial role in the design of high performance engineering applications such as aircraft and naval vessels. Current turbulence models are known to give erroneous predictions of onset of separation and reattachment lengths. Recent years have seen an increase in availability of high fidelity data sets; and thus, data driven modelling is now being tested as a potential tool to improve turbulence closure models. In line with this goal, the present study aims to evaluate the machine learning as a means to augment turbulence modelling.

Empirical data is obtained for a series of increasingly high bumps by Large Eddy Simulation. A patch of high turbulent kinetic energy forms in the lee of the bump and extends into the wake. It originates near the surface and has a significant influence on flow development. The highest bumps create a small separation bubble. Over the bump the log-law is absent, evidencing strong disequilibrium. The dataset is created to be used in data-driven modelling.

An optimization method is used to extract fields of variables that are used in turbulence closure models. From this, it is shown how these models fail because they predict near-wall eddy viscosity erroneously. Machine learning is used to generalize the optimized field variables such that existing turbulence models can produce more accurate results on different test cases. It is shown that these machine learning augmented closure models result in a modest improvement in test cases.

CHAPTER 1. INTRODUCTION

Turbulence is ubiquitous in nature, with a large number of flows in the universe, ranging from the every day stirring of coffee to the accretion disk around a black hole, being turbulent. Studying turbulence, although not easy by any means, is hence an important step in understanding the world around us. Even more difficult is building the capability of assimilating this knowledge to make predictions in "real world" scenarios.

An explicit definition of turbulence is difficult to state, however turbulent flows are known to have the following characteristics (Tennekes et al., 1972)

(i) Randomness

(ii) Increased Diffusivity

(iii) Three Dimensional fluctuations

(iv) Increased Dissipation

Turbulence is a characteristic of the flow, not of the fluid, and is governed by the Naiver Stokes equations (1.1). Note that this and all the following equations are valid for incompressible flows only. The non-linear nature of the governing PDEs can result in turbulence and makes finding theoretical solutions to these equations almost intractable (except in special cases) and no general solutions to them are known.

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} \tag{1.1}$$

However with the advent of modern computers, numerical solution of complex fluid phenomena (for a set initial and boundary conditions) has become possible. The problem now lies in having

sufficient spatial and temporal resolution to obtain accurate solutions. Turbulent flows have multiple scales (length and time) present simultaneously,and a higher resolution means resolving smaller scales which in turn requires more computational time.

1.1 Simulation of turbulent flows

1.1.1 Direct Numerical Simulations

Direct Numerical Simulation (DNS) involves solving the Navier Stokes equation on a grid which is fine enough to resolve the smallest of length scales and has a time step sufficiently small to capture the fastest fluctuations at these length scales. With such resolution, direct numerical simulations are considered to essentially be exact, with very small numerical errors. Consequently, they are the most expensive kind of simulations.

1.1.2 Large Eddy Simulations

A less expensive kind of simulation is Large Eddy Simulation (LES) in which the smallest of scales are not simulated but "modelled". The large energy and momentum carrying structures are simulated and thus these simulations are still fairly accurate. Since large eddy simulations only model the smaller scales, which tend to be homogeneous and less affected by boundary condition (Piomelli, 2014), the models are relatively easier and thus more accurate.

The resolution of the simulation depends on grid and time step refinement. Since LES is only supposed to capture the largest of "eddies" the grid is only refined to capture the energy containing eddies. This means that the grid is only capable of capturing the larger fluctuations. Thus the field that the grid captures can be regarded as a filtered field in which the smallest of the fluctuations have been filtered out. An evolution equation for such a filtered field can be derived by filtering the equations of motion (1.1) to get 1.2.

$$\frac{\partial \overline{u_i}}{\partial t} + \overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j} = -\frac{1}{\rho}\frac{\partial \bar{p}}{\partial x_i} + \nu\frac{\partial^2 \overline{u_i}}{\partial x_j \partial x_j} - \frac{\partial \tau_{ij}}{\partial x_j} \tag{1.2}$$

where the overline represents the filtered field and $\tau_{ij} = \overline{u_i u_j} - \overline{u_i}\,\overline{u_j}$ are called the "subgrid-scale (SGS)" stresses as they signify the effect of the unresolved field (filtered out field) on the resolved field (filtered field). These SGS stresses need to be modeled and there exist multiple ways to do so. The method used in the present work to model them will be discussed in later chapters.

DNS and LES simulations are inherently time dependent solutions (even for statically steady problems) and the solutions need to be averaged (in time or ensemble averaged for statically unsteady problems) for them to be of any use to the engineering community. This is because a majority of problems require a single valued answer (averaged): for example an aircraft designer would be more interested in knowing a single value of the drag coefficient (at a particular Reynolds number, Mach number and angle of attack) than a time dependent series of drag coefficient values which vary over the course of the simulation (which is indeed the case with turbulent flow since the velocity gradients and hence the drag will continuously change, about a certain mean, as the flow fluctuates over the airfoil!).

1.1.3 Reynolds Averaged Naiver Stokes Simulations

The most prevalent way to simulate turbulent flows is using the Reynolds averaged Naiver Stokes (RANS) equations. These computations attempt to predict statistics directly, invoking semi-empirical models. Hence, these simulations are the least accurate of the three but the most computationally viable. RANS simulations can be steady or unsteady depending on the problem. For a statistically steady state problem, RANS equations can be solved without having a dependence on time.

Although computers have come a long way it is still impractical to use DNS and to a large extent LES for large scale flows and thus although inaccurate, we are forced to use RANS simulations. Since these are the only kind of simulations that are computationally affordable, in most cases, a large amount of effort is dedicated to increasing the accuracy of the *models* used in RANS equations.

RANS equations (1.3) are derived by taking the average of the Naiver Stokes equations. This operation results in the appearance of an extra term $-\overline{u_i^{'} u_j^{'}}$ where $u^{'}$ denote the fluctuations and

the over-line denotes an average ($u^{'} = u - \overline{u}$). This term is often referred to as the Reynolds stress which is the averaged effect of turbulent convection (Durbin and Reif, 2011).

$$\frac{\partial \overline{u_i}}{\partial t} + \overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j} = -\frac{1}{\rho}\frac{\partial \overline{p}}{\partial x_i} + \nu\frac{\partial^2 \overline{u_i}}{\partial x_j \partial x_j} - \frac{\partial \overline{u_i^{'} u_j^{'}}}{\partial x_j} \tag{1.3}$$

The emergence of the Reynolds stress results in more unknowns than equations, which is known as the closure problem. Reynolds stress, thus, has to be modeled. The basic idea of RANS equations is to solve for the averaged field (which we would get theoretically by taking the average of DNS) without having access to the observations (time dependent data). This in essence is the turbulence closure problem. There are many different ways to estimate the Reynolds stresses and some of them are discussed in the next chapter.

1.2 Overview

The thesis has the following parts:

(i) A brief description of closure models and description of methodology.

(ii) Description of LES simulations and conversion of the data into "usable form" using an optimization procedure.

(iii) Using Machine learning to improve closure model by utilizing the optimized data.

CHAPTER 2. RANS MODELS AND METHODOLOGY

In order to solve the RANS equations we need to predict the value of the Reynolds stresses. Over the course of time this has been done in many different ways. We will focus on what are know as single point closures, which means that we try to estimate the Reynolds stress at a point using information (velocity gradients, etc) only at that point. In particular we will focus on a sub category, two equation single point closure models. These models make an assumption (not necessarily a correct one!) that the Reynolds stress can be related to the mean flow using a Newtonian constitutive equation (2.1) with an eddy viscosity, ν_T.

$$\overline{u_i^{'}u_j^{'}} = -2\nu_T S_{ij} + \frac{2}{3}\delta_{ij}k \tag{2.1}$$

In a Newtonian fluid, the shear stress is linearly dependent on the velocity gradient; this relation (2.1) assumes that in a turbulent flow, the Reynolds stress is linearly dependent on the velocity gradient . In this "Boussinesq approximation", the coefficient of proportionality is the eddy viscosity. This is why these models are also known as linear eddy viscosity models (LEVMs). Note: In the constitutive equation we generally only concern ourselves with the $2\nu_T S_{ij}$ part and group the $2/3\delta_{ij}k$ part with the pressure in the momentum equation and solve for a "modified pressure". The hurdle in predicting the Reynolds stress is predicting ν_T. This problem can again be solved in a variety of different ways.

Since the eddy viscosity has the units (m^2/s) it can be written as the product of a length (l^*), time (T^*) and velocity scale (u^*) as in equation 2.2. In the models that will be discussed here, u^*, T^* and l^* are related to other turbulence variables like the turbulent kinetic energy (TKE), dissipation, etc. (Pope, 2001).

$$\nu_T = u^* l^* = u^{*2} T^* \tag{2.2}$$

The two most popular, two equation models to obtain the eddy viscosity are by using either the $k-\epsilon$ or the $k-\omega$ models.

2.1 Two equation closure models

2.1.1 k-ϵ model

The standard $k-\epsilon$ model relates (2.3) the eddy viscosity to the turbulent kinetic energy (k) and the rate of dissipation(ϵ) via

$$\nu_t = C_\mu \frac{k^2}{\epsilon} \tag{2.3}$$

with $C_\mu = 0.09$. The relation can either be derived using simple dimensional analysis or by rationalizing that at high Reynolds number, the rate of dissipation and production are of similar order of magnitude (Durbin and Reif, 2011). $k^{1/2}$ is used as the velocity scale and the length scale is given by $k^{3/2}/\epsilon$. The actual TKE equation is given by 2.4

$$\frac{Dk}{Dt} = -\nabla \cdot T + P - \epsilon \tag{2.4}$$

where $T_i = \frac{1}{2}\overline{u_i' u_j' u_j'} + \overline{u_i' p'}/\rho - 2\nu \overline{u_i' S_{ij}'}$. This term is unclosed and is thus replaced by the gradient transport model and we get 2.5.

k and ϵ are determined by solving the following transport equations.

$$\frac{Dk}{Dt} = 2\nu_t |S|^2 - \epsilon + \nabla \cdot (\nu + \nu_t)\nabla k \tag{2.5}$$

$$\frac{D\epsilon}{Dt} = 2c_{\epsilon 1}\frac{\epsilon}{k}|S|^2 - c_{\epsilon 2}\frac{\epsilon^2}{k} + \nabla \cdot (\nu + \sigma_\epsilon \nu_t)\nabla \epsilon \tag{2.6}$$

with some modifications near to walls.

It can be argued because of its origins, the k equation in the model is exact. On the other hand, the transport equation for ϵ is almost purely an "invention" (Davidson, 2015) with tunable empirical constants ($C_{\epsilon 1}, C_{\epsilon 2}$ and σ_ϵ). These empirical constants are tuned using standard flows

and are not universally applicable. The ϵ equation needs near wall modifications to prevent a singularity and the eddy viscosity needs to be dampened in the vicinity of a wall to get the correct eddy viscosity. These and other shortcomings with this model makes it difficult to recommend its usage near wall. Note there are several ways to solve RANS equations with the $k-\epsilon$ model and curcumvent the near wall problems by using either wall functions or two layer models. Another way to fix the near wall issue is to use a different model entirely which has better near wall performance.

2.1.2 k-ω model

Similar to the $k-\epsilon$ model, the $k-\omega$ model equates (2.7) the eddy viscosity to the TKE and specific rate of dissipation (ω).

$$\nu_t = \frac{k}{\omega} \tag{2.7}$$

k and ω are determined by solving the following transport equations.

$$\frac{Dk}{Dt} = 2\nu_t|S|^2 - C_\mu k\omega + \nabla\cdot(\nu+\nu_t)\nabla k \tag{2.8}$$

$$\frac{D\omega}{Dt} = 2c_{\omega1}|S|^2 - c_{\omega2}\omega^2 + \nabla\cdot(\nu+\sigma_\omega k/\omega)\nabla\omega + (\sigma_d/\omega)\nabla k\cdot\nabla\omega \tag{2.9}$$

These can be solved to the wall, with suitable boundary conditions. $k^{1/2}$ is again used as the velocity scale and "specific dissipation rate": ω is used as an inverse time scale. The k equation in the k-ϵ model is only modified by substituting $\epsilon = C_\mu k\omega$ (Wilcox, 2008). The ω equation is again ad-hoc. The last term in the ω equation ($(\sigma_d/\omega)\nabla k\cdot\nabla\omega$) is known as the cross-diffusion term and is helpful in reducing the free stream sensitivity of the model (Wilcox, 2008). Several variations of the model exists and the one used here is one of the later ones ($k-\omega$ 2006). The k and ω equations are coupled and since the ω equation "behaves" differently than the ϵ equation in the $k-\epsilon$ model, the k equations also yields different k profiles with the k in the $k-\omega$ model not having the same asymptotic behavior as TKE in DNS data. Therefore k in the $k-\omega$ model is not really TKE

(Durbin and Reif, 2011). Nevertheless, the near wall ν_T predicted by this model better agrees with DNS data (Durbin and Reif, 2011) compared to the $k - \epsilon$ model.

2.2 Methodology

In the first step, LES was used to generate high resolution data for the flow field and was regarded as the "ground truth". Several parameters were used to ensure the accuracy of the data and will be discussed in detail in the upcoming chapter. Second, an optimization problem was solved to convert this data into a model usable form. One of the model coefficients ($c_{\omega 1}$ or $c_{\omega 2}$) was optimized for a flow configuration using the LES data. This meant that the coefficient was no longer constant through out the field and became a function of the position. Lastly, machine learning methods were used to form a correlation between the coefficient field and flow parameters. Utmost care was taken to ensure that the flow parameters (features) were rotation and translation invariant, and non dimensional. This step was to ensure that the correlation could be ported to other flow configurations making this setup comparable to a standard turbulence model.

The first two steps are discussed in chapter 3 and the third step is discussed in chapter 4.

CHAPTER 3. LES AND FIELD INVERSION

Adapted from Matai R, Durbin PA (2018) LES of turbulent flow over a parametric set of bumps. to be submitted

Abstract

Turbulent flow over a series of increasingly high, two dimensional bumps is studied by well resolved large-eddy simulation. Predictions of mean flow and Reynolds stress for the lowest bump are in good agreement with experimental data. A grid sensitivity study is carried out for the highest bump to characterize the accuracy of the simulations. The flow encounters a favourable pressure gradient over the windward side of the bump, but does not re-laminarize as is evident from near wall fluctuations. A patch of high turbulent kinetic energy forms in the lee of the bump and extends into the wake. It originates near the surface, before flow separation and has a significant influence on flow development. The highest bumps create a small separation bubble, whereas flow over the lowest bump does not separate. The log-law is absent over the entire bump, evidencing strong disequilibrium. The peaks present in the Reynolds stresses increase in magnitude but do not shift in location (y_+ units) as the bump height increases.Turbulent kinetic energy production spreads out away from the wall as the flow progresses over the adverse pressure gradient region of the bump.

The dataset was created to be used in data-driven modeling. Present RANS models are unable to correctly predict velocity profiles for this set of bumps. It is shown that these models fail because they predict near-wall eddy viscosity erroneously. The discrepancies extend further away from the wall in the adverse pressure gradient and recovery regions than in the favourable pressure gradient region.

3.1 Introduction

Predicting the development of boundary layers in adverse pressure gradient remains a driving force for improving various types of turbulence models (Slotnick et al., 2014). Some Reynolds averaged models predict premature onset of separation; some under predict the rate of recovery after the end of an adverse pressure gradient region (Durbin, 2018). In Hybrid RANS-LES, sometimes eddies develop too slowly, resulting in incorrect predictions. For these reasons, adverse pressure gradient test cases have become standard for model development and validation. Flow over bumps have become a source of data for such studies (Seifert and Pack, 2002; Breuer et al., 2009). While the motive is to investigate adverse pressure gradient (APG), the upstream side of the bump is exposed to favorable pressure gradient (FPG). That does not mitigate the usefulness of the data, but now the data are for a particular geometry, and one is looking more broadly at pressure gradient effects. The bump geometry is similar to geometries encountered in many engineering problems, such as the upper side of an airfoil, or a duct constriction, and is therefore an interesting study.

The effect of pressure gradients, with and without curvature, have been studied extensively. Experiments and simulations have explored different aspects of pressure gradient flows such as internal layer formation, relaminarization, etc. Blackwelder and Kovasznay (1972) reported that a strong favourable pressure gradient resulted in the disappearance of the 'law of the wall' and caused a decrease in the intermittency within the boundary layer. (This intermittency is different from intermittent flow reversal, discussed later.) It was also observed that a more 'laminar-like' state had lower turbulent kinetic energy (TKE) and $-\overline{uv}$ in the inner layer, and larger values in the outer layer, than the original turbulent state. It will be seen in the present case that the TKE is higher than the flat plate regions, over most of the accelerating region. A self similar, APG turbulent boundary layer over a flat plate was simulated by Kitsios et al. (2016). Reynolds stresses and turbulent kinetic energy (TKE) were shown to develop a second outer peak. The outer production peak was associated with the additional mean shear caused by the pressure gradient.

Another interesting characteristic of some wall-bounded, pressure gradient flows is the formation of internal layers.

In a traditional equilibrium boundary layer, the mean velocity is in equilibrium with the surface. An internal layer forms if the flow encounters a change in surface geometry. The profile is modified, starting at the wall and growing into the flow. This is similar to the formation of a boundary layer when a uniform velocity encounters a surface (and is thus subjected to a different boundary condition). An internal layer is signified by the presence of knee points in the Reynolds stress profiles (points of inflection in the velocity profile). Tsuji and Morikawa (1976) experimentally studied the effect of alternating pressure gradient on a flat plate boundary layer, observing the formation of an internal layer when a positive pressure gradient followed a negative pressure gradient. An analogy was drawn between the free stream pressure gradient and surface curvature. Thus, a change in the sign of pressure gradient would be equivalent to a sudden change in surface curvature. In these experiments, Reynolds stresses were conserved along streamlines beyond the knee points and changed in the inner part of the layer. The change near to the wall was attributed to the abrupt change of pressure gradient.

A numerical and experimental study by Spalart and Watmuff (1993) of turbulent boundary layers with pressure gradients revealed that the buffer and lower end of the log layer shift up in FPG and down in APG. Bandyopadhyay and Ahmed (1993) also noticed a departure from the log-law, and observed that skin friction recovered faster following convex-to-concave geometry than following concave-to-convex. Alving et al. (1990) reported a slow recovery of the skin friction and Reynolds stresses. They reported that even after "30 local boundary layer thicknesses, little relaxation actually occured".

Gillis and Johnston (1983) performed experiments to study the effect of strong stabilizing curvature ($\delta_{ref}/R = 0.05$ and $\delta_{ref}/R = 0.10$) on boundary layer turbulence where δ_{ref} was the reference boundary layer thickness. They noted that the flow recovered very slowly on a flat section, after the curvature. Curvature reduced the size of large scale eddies, which reduced the influence of upstream conditions, and the turbulent structures became more dependent on R. A shortened logarithmic region was observed, ending near $y_+ = 100$.

Jesus et al. (2013) used wall-resolved LES to simulate turbulent flow in a channel with two different bump geometries, finding excellent agreement with DNS. Flow was simulated at $Re_\tau = 617$ and 950. The study also showcased the poor performance of various turbulence models, two equation and Reynolds stress transport, over this geometry. The models failed to correctly predict the skin friction in the APG and the Reynolds stress transport model overestimated normal Reynolds stresses near the bottom wall. LES on a coarse mesh was shown to give satisfactory skin friction values. In a similar LES study, Schiavo et al. (2015) observed that production peaks of the Reynolds stresses shifted away from the wall at the separation point. Kuban et al. (2012) studied the effects of grid resolution and subgrid scale (SGS) models on flow solutions, and found that streamwise resolution was more important than spanwise and wall normal resolution, for accurate solutions.

Most data sets are for a single geometry. This creates uncertainty whether predictions are peculiar to that particular case. Here, we address this concern about ideosyncrasy by considering a family of bumps. (Mollicone et al. (2017) also studied a family of bumps but all their bumps had separated flow.) By considering a parametric family of bumps, systematic behaviors can be explored. Aspects of data that are representative for the class of flows can be distinguished from specific data for a single geometry.

For the present study, the circular arc bump of Webster et al. (1996) is adopted; see fig. 3.1. The Webster *et al.* paper was presented, simply, as a database contribution. In reproducing the experiment by large eddy simulation (LES), we noticed some intriguing features which are discussed herein. A parametric series is then created by increasing the height of the bump, keeping the length fixed. The baseline case has $L = 254mm$ and $h = 20.1mm$. Convex fillets are added before and after the bump to create a total length of $C = 305mm$. The series is parameterized by h/C starting with this ratio, 20/305, and increasing up to 42/305. The flow has separated by that value. Data sets are generated by well resolved LES and a grid refinement study is conducted to ensure the accuracy of the results. The baseline geometry was previously simulated at a lower Reynolds number by Wu and Squires (1998b), who were able to reproduce the experimental measurements via LES.

The Webster et al. (1996) data have been used to test models (Parneix et al., 1998; Wu and Squires, 1998a). To that end it would be valuable to have data on turbulence model variables (e.g., ν_T, k, ω). However, these variables are defined by their role in the model and cannot be measured directly as physical data. Even the variable k, which is named turbulent kinetic energy, does not behave as the measurable turbulent kinetic energy. We use the innovative method of Duraisamy et al. (2015); Parish and Duraisamy (2016a) to extract data on model variables by solving an optimization problem. Preliminary results for our data are contained in Singh et al. (2017).

3.2 Simulation overview

The flow over a parameterized set of bumps was simulated by wall resolved LES in the finite volume code OpenFOAM (Weller et al., 1998). The code uses a co-located grid arrangement with support for unstructured grids. The simulations were carried out using the unsteady and incompressible flow solver "pisoFoam" which uses the PISO algorithm to solve the Navier-Stokes equations. The second order, backward, implicit time advancement scheme, was selected. 'Gauss linear' discretization was used for gradient and divergence terms, where 'linear' stands for the linear interpolation scheme (central differencing, second order accurate) to interpolate values from cell centers to face centers and 'Gauss' specifies Gaussian integration. Laplacians were also discretized with the 'Gauss' scheme and surface normal gradients were computed with the 'corrected' scheme which is also second order accurate. The pressure and momentum equations were solved by preconditioned (bi-)conjugate gradient. Diagonal incomplete-Cholesky factorization (for symmetric, pressure equation) and incomplete-LU factorization (for asymmetric, momentum equation) was used for preconditioning. The pressure and momentum equations correctors were solved two times in each step. The subgrid stresses were represented by the dynamic Smagorinsky subgrid model as in equation 3.1 (Lilly, 1992).

$$\tau_{ij} - \frac{1}{3}\tau_{kk} = 2C\Delta^2|\overline{S}|\overline{S_{ij}} \tag{3.1}$$

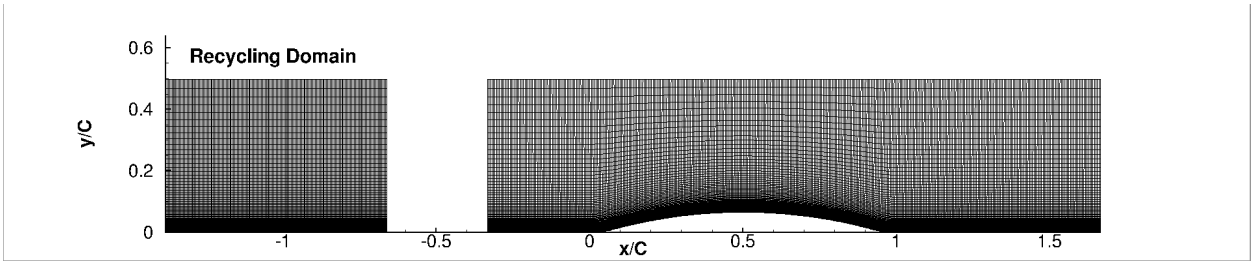


Figure 3.1 Mesh used for the LES simulation, shown for the h20 reference bump. Note the disconnected recycling domain for generating the inlet flow ($C = 0.305$m).

and C is calculated at each computational node and averaged locally:

$$C = \frac{1}{2}\frac{L_{ij}M_{ij}}{M_{ij}^2} \tag{3.2}$$

where,

$$L_{ij} - \frac{1}{3}L_{kk} = 2CM_{ij} \tag{3.3}$$

and,

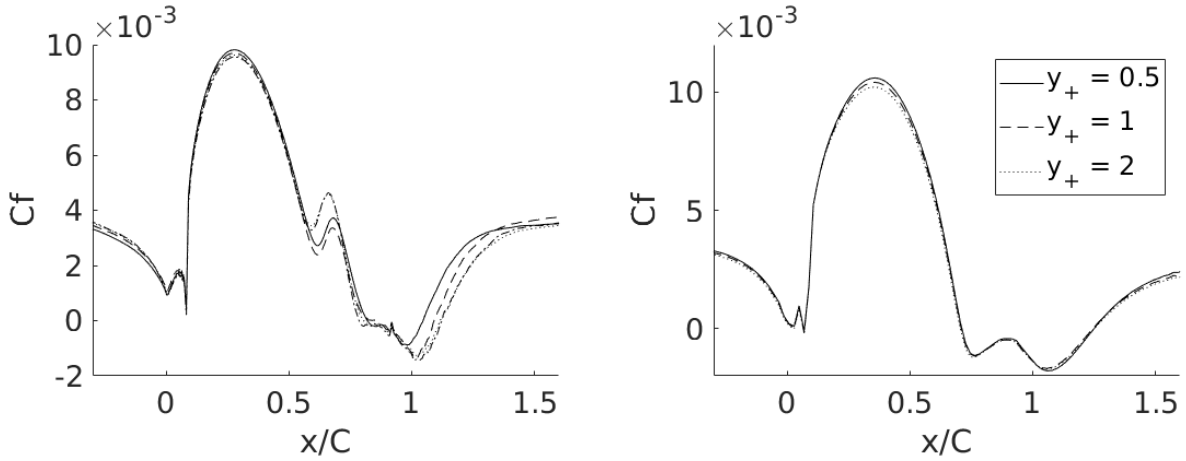
$$M_{ij} = \widehat{\Delta^2}|\widehat{\overline{S}}|\widehat{\overline{S_{ij}}} - \Delta^2|\overline{\widehat{S}}|\overline{\widehat{S_{ij}}} \tag{3.4}$$

A time dependent, turbulent inflow was generated by a separate, concurrent flat plate boundary layer simulation as can be seen in Fig. 3.1. The disconnected upstream domain recycles and rescales the flow (Arolla, 2016) to generate fully developed inflow conditions for the downstream domain. The inlet momentum thickness was $\Theta = 3.6mm$; hence, for the baseline case $\Theta/h = 0.18$ and $\delta_{99}/h = 1.5$. This same thickness was used for the higher bumps. The time step, $dt = 0.009\theta/U_{ref}$ was used (max CFL number < 0.7), where U_{ref} is the inlet free stream velocity.

The inlet to the bump is located at $x/C = -1/3$ with an inlet Reynolds number of 2,500 based on inlet free stream velocity and momentum thickness ($= C/82$). The bump starts at $x/C = 0$ and ends at $x/C = 1$ and the width of the domain is $0.22C$. The top boundary is zero normal-gradient for the velocity and pressure and the outlet condition is zero normal-gradient for the velocity and zero pressure (reference pressure).

Table 3.1 Grid characteristics of the bump simulation at $Re_\theta = 2500$

#	Δy_+	$\Delta x_+(APG, FPG/ZPG)$	Δz_+	Number of cells	N_x, N_y, N_z
Mesh 1	<0.5	41	17	7 million	584, 109, 110
Mesh 2	<0.33	23,41	17	13 million	767, 159, 110
Mesh 3	<0.33	23,41	11	18 million	767, 159, 154
Mesh 4	<0.33	18,41	9	27 million	876, 159, 199



(a) LES (b) RANS

Figure 3.2 Grid Convergence for highest bump (h42), LES: —— Mesh 1, - - - Mesh 2, ····· Mesh 3, — - — Mesh 4. RANS:—— $y_+ = 0.5$, - - - $y_+ = 1$, ····· $y_+ = 2$. Note y_+ is at the inlet.

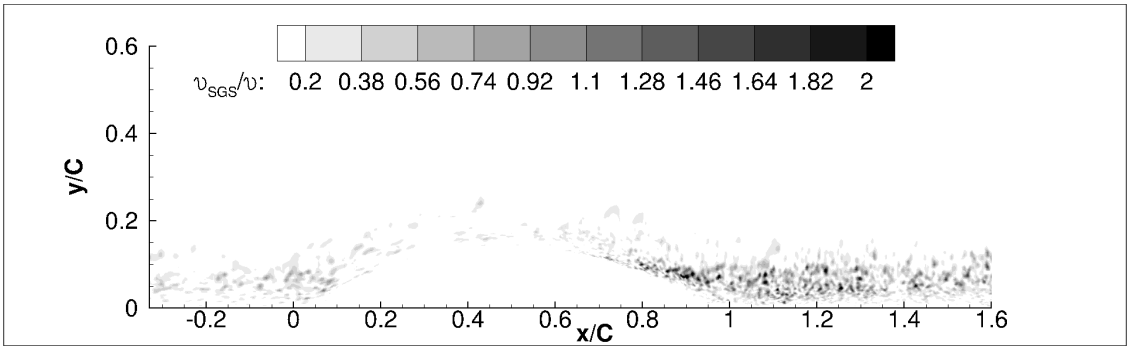


Figure 3.3 Ratio of subgrid viscosity to kinematic viscosity for highest bump (h42).

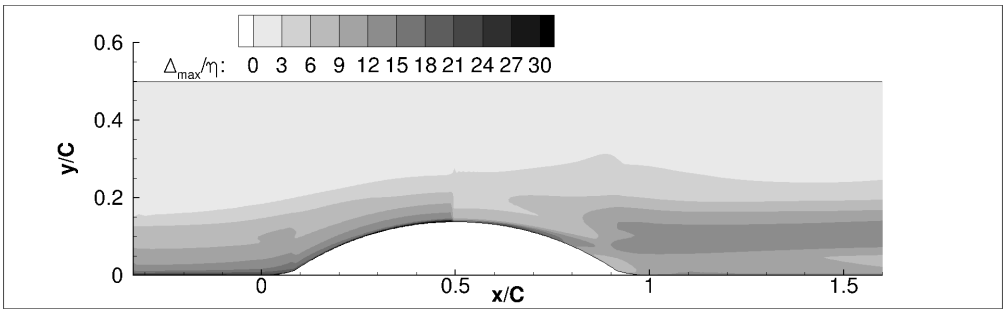
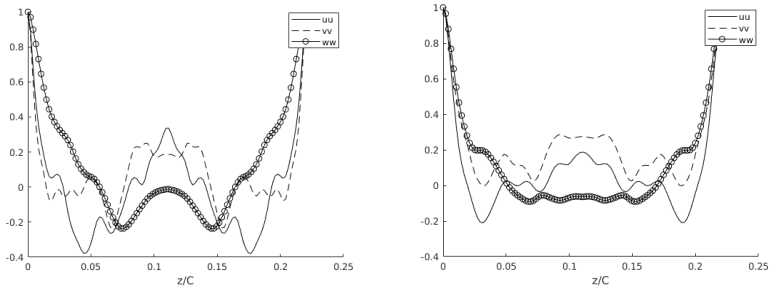


Figure 3.4 Ratio of maximum grid spacing to Kolmogorov length scale for highest bump (h42).



(a) $y/\delta = 0.2$ (b) $y/\delta = 0.4$

Figure 3.5 Spanwise correlation function for —— $\overline{uu}$, - - - $\overline{vv}$ and ∘ $\overline{ww}$ at $x/C = -0.26$.

3.2.1 Verification

As the highest bump has the largest separation and requires the finest grid, a sensitivity study was carried out on this bump. This same resolution was then used for all the bumps. Table 4.1 gives the grid characteristics for the different resolutions that were tested and fig. 3.2 shows the skin friction for these meshes. The grid spacing is in plus units, using the friction velocity at the inlet. LES is inherently grid dependent. Hence, the grid dependence in fig. 3.2 can be regarded as a measure of experimental accuracy. The accuracy is very good, and even better for the lower bumps; indeed, the reference bump is virtually grid independent.

It was observed that resolution in the span and streamwise directions was more important for accuracy than was the wall normal direction. The streamwise resolution was most important in the APG region. As seen in fig. 3.2, the upstream region is virtually grid independent. Hence,

streamwise refinements were made after the crest of the bump; before the crest, all meshes have the same streamwise resolution.

Mesh3 and mesh4 had identical skin friction and velocity profiles, indicating that mesh3 had sufficient resolution and, thus, was chosen for all the simulations. Flow over these bumps was also computed by RANS with the $k-\omega$ closure model. To ensure the RANS solution did not depend on the near wall resolution, the y_+ dependency was studied (fig. 3.2). It was observed that three different values of y_+ had minimal effect on the flow solution and the mesh with inlet $y_+ = 0.5$ was used in all subsequent RANS simulations.

The ratio of the subgrid viscosity to molecular viscosity is a standard metric for LES resolution. Fig. 3.3 shows that the ratio is less than 2 over most of the domain, which indicates very good LES resolution (Durbin and Reif, 2011). Fig. 3.3 shows that the viscosity ratio continuously increases as the flow progresses over the bump, is the highest over the adverse pressure gradient region, and then decreases over the recovery region. The higher ratio in the adverse pressure gradient (APG) region suggests a higher resolution is needed to resolve the smaller scales present there. Another metric to assess resolution of the simulations is the ratio of grid spacing to the Kolmogorov length scale. The ratio was less than 30 for the highest bump as shown in fig. 3.4; this also indicates good resolution. Note that these metrics were evaluated on the highest bump using mesh3 and grid spacing is the length of the largest side of a cell.

Figure 3.11, below, shows that the domain width is sufficient to capture several high and low speed streaks, indicating that the domain width was sufficient in the z-direction. Spanwise correlations are not as definitive. A converged correlation was not computed. Fig. 3.5 shows a short time average spanwise correlation for u, v and w at $x/C = -0.26$, just in front of the highest bump. It is indicative that the width is sufficient. Reynolds averaged statistics were computed with a wider domain and found to be virtually unchanged.

In order to verify the budget analysis, a zero pressure gradient flat plate boundary layer was simulated and the budgets were compared to the DNS by Spalart (1988). The domain size was

Table 3.2 Grid characteristics of flat plate boundary layer simulation to validate budgets.

Re_θ	Δy_+	Δx_+	Δz_+	Number of cells	N_x, N_y, N_z
1410	<0.5	18	5.5	3.33 million	222, 75, 200

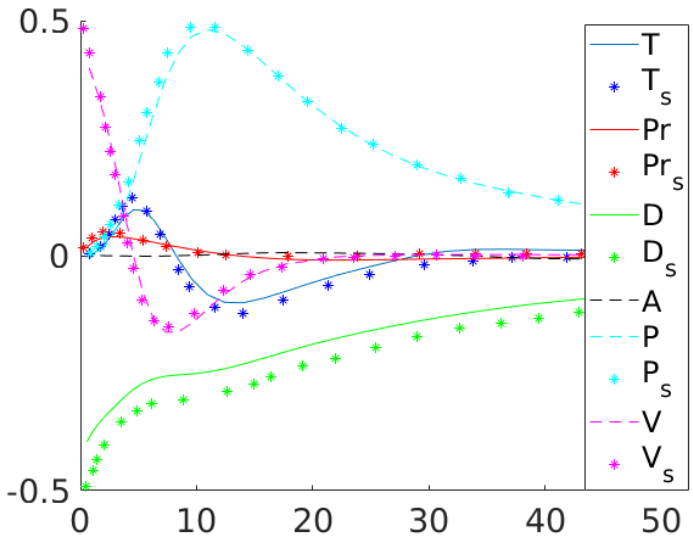


Figure 3.6 Comparison of budget of TKE with DNS. T: Turbulent transport, Pr: Pressure transport, D: Resolved Dissipation, A: Advection, P: Production, V: Viscous transport, s: subscript for Spalart's DNS data.

equal to the recycling zone shown in fig. 3.1 and the same recycling method was used to obtain $Re_\theta = 1,410$. The grid characteristics are shown in table 4.2.

The TKE budgets from the LES simulations matched reasonably well with the DNS data (fig. 3.6). It can be seen that the resolved dissipation is less than the DNS; this is expected because of the presence of subgrid dissipation in the LES. However, it can be seen that a lower resolution only slightly shifts the peaks of the budget, and still delivers accurate results. Note that the ZPG flat plate was only simulated to test the budget code and therefore simulated on a finer grid than the bump.

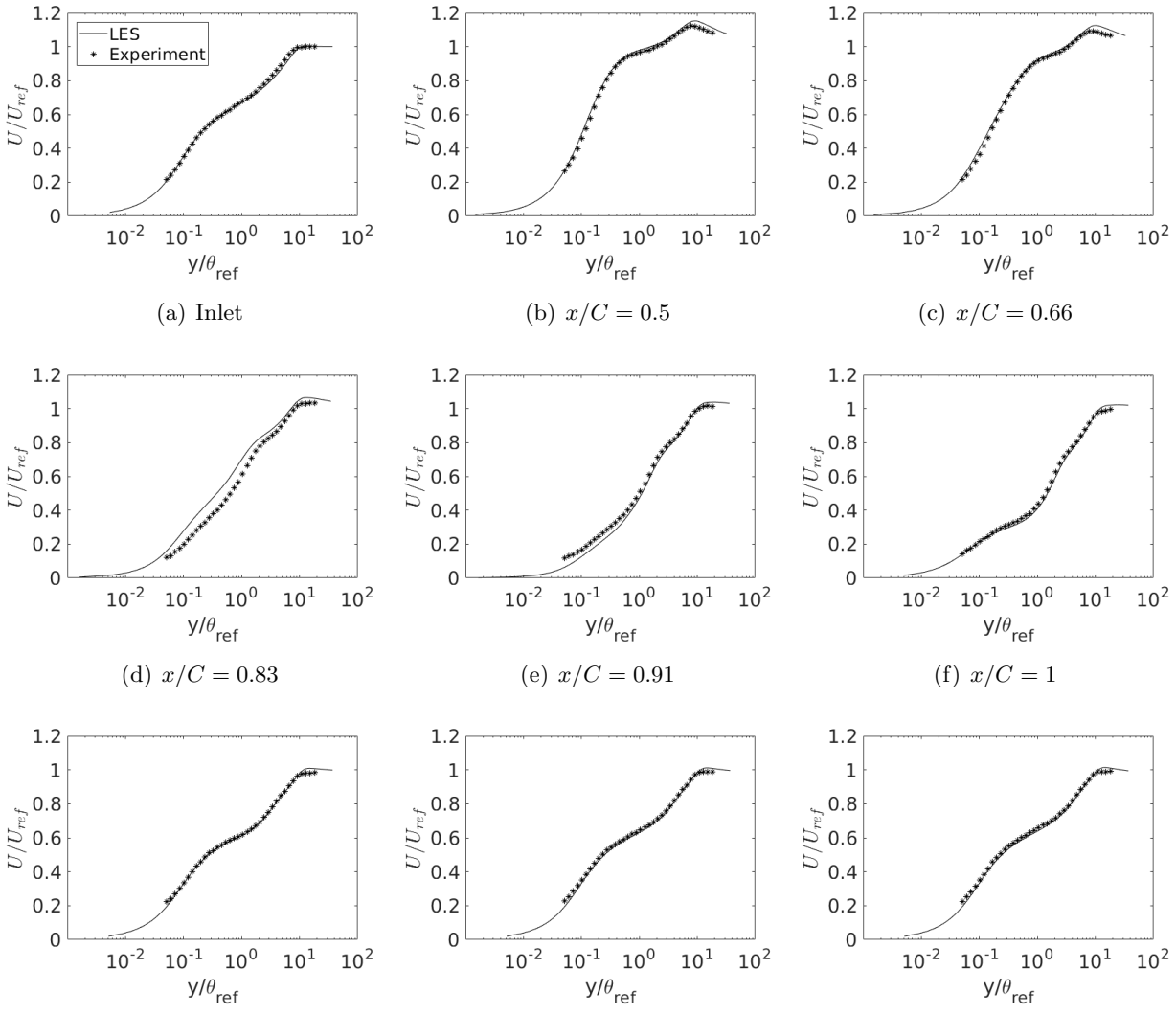
3.2.2 Validation

Time and span averaged velocity profiles were compared with experimental data (available for the lowest bump) from Webster et al. (1996) at nine locations (fig. 3.7). Components of Reynolds stresses, $\overline{uu}$ (fig. 3.8), $\overline{vv}$, $\overline{ww}$ and $\overline{uv}$ (fig. 3.9), were also compared. Velocity is scaled on the inlet free stream velocity ($16.77m/s$), and the inlet momentum thickness ($0.0037m$) is the length scale.

The LES profiles of mean velocity in the x-direction match reasonably well with the experimental data, except at $x/C = 0.83$; even there, the trend of the LES profile does follow the experimental data. Wu and Squires (1998b) (LES at $Re_\theta = 1,500$) reported similar deviation of the LES profile at $x/C = 0.91$ and attributed it to strong intermittent flow reversal at the wall. This could be the case here as well, since the intermittent reversal starts around $x/C = 0.79$ at the present Reynolds number. Over the bump, fig. 3.7 shows large deviations from the log law; this can be seen clearly in a later section where velocity is with normalized by viscous units. The experimentalists noted this, too, and used an oil drop method to infer skin friction, because of the strong departure from equilibrium. The flow is close to recovering to a flat plate solution by $33\%C$ as seen by the velocity profiles at $x/C = 1.33$, 1.5 and 1.63, which are fairly similar to the inlet mean velocity profile. Although the flow is close to equilibrium by $x/C = 1.33$, it can be seen that the flat plate, inlet profile is not recovered until the end of the domain (this will be clearer in the following section when the profiles are plotted in viscous units). Thus even after $5\delta_{inlet}$ the flow does not recover. This slow recovery is in line with observations by Alving et al. (1990).

The $\overline{uu}$ profiles also match well with the x-wire experimental data, where they are available. However, the LES shows structure close to the wall that could not be measured by the x-wire. The experimenters measured some data with a single wire, but they are not quantitatively accurate. The single wire data are included to show that the near-wall behaviour seen in the LES was also present in the lab experiment.

The peak that appears beyond $y/\theta_{ref} = 1$, at the end of the bump ($x/C = 1$), is seen in the x-wire data. This peak flattens as the flow recovers after the bump. It will be seen in contour plots

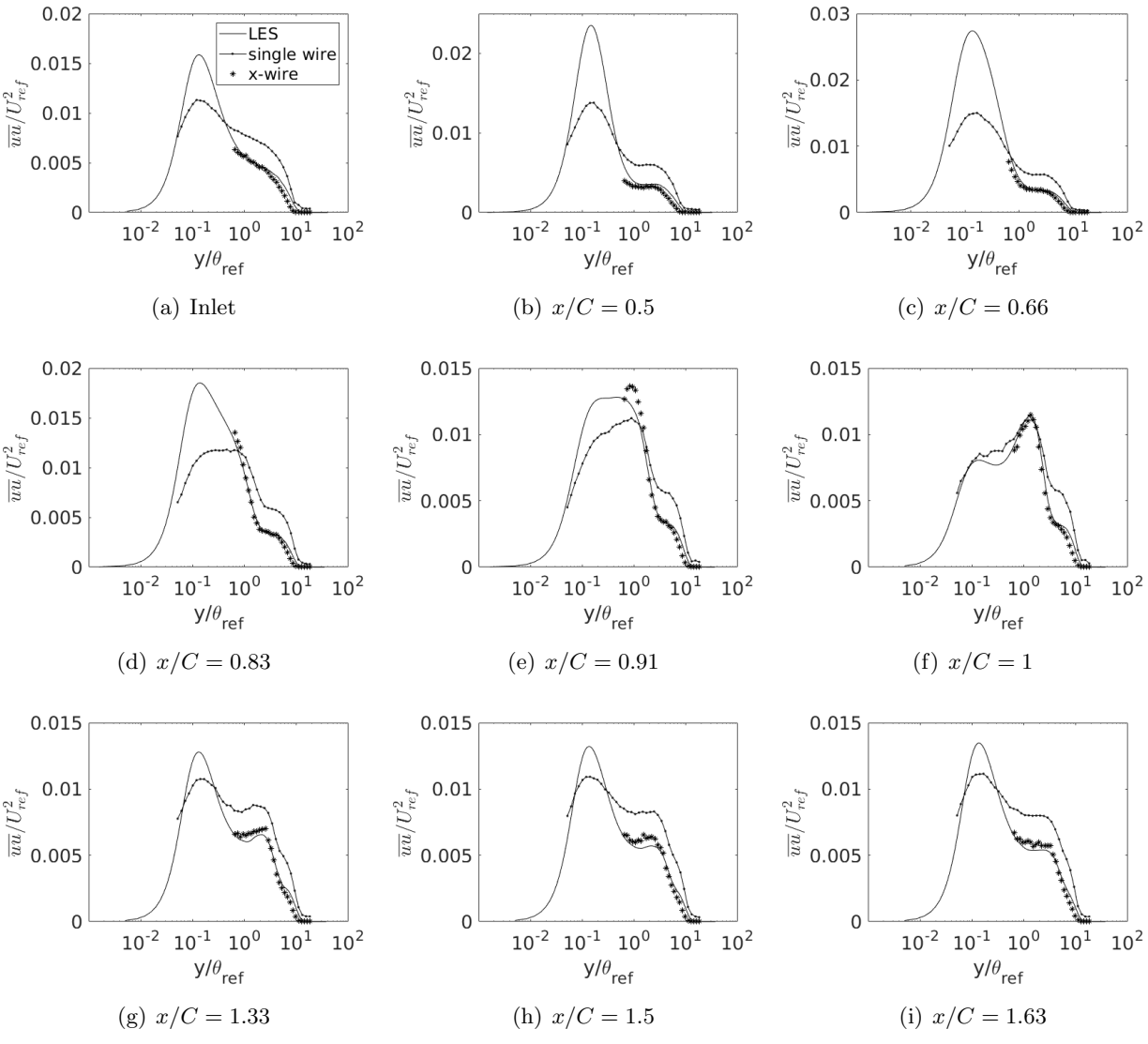


(a) Inlet (b) $x/C = 0.5$ (c) $x/C = 0.66$

(d) $x/C = 0.83$ (e) $x/C = 0.91$ (f) $x/C = 1$

(g) $x/C = 1.33$ (h) $x/C = 1.5$ (i) $x/C = 1.63$

Figure 3.7 Time and span averaged x-velocity at various x locations. —— LES, $*$ experiment.



(a) Inlet (b) $x/C = 0.5$ (c) $x/C = 0.66$

(d) $x/C = 0.83$ (e) $x/C = 0.91$ (f) $x/C = 1$

(g) $x/C = 1.33$ (h) $x/C = 1.5$ (i) $x/C = 1.63$

Figure 3.8 Time and span averaged $\overline{uu}$ at various x locations. —— LES, —•— single wire, $*$ x-wire.

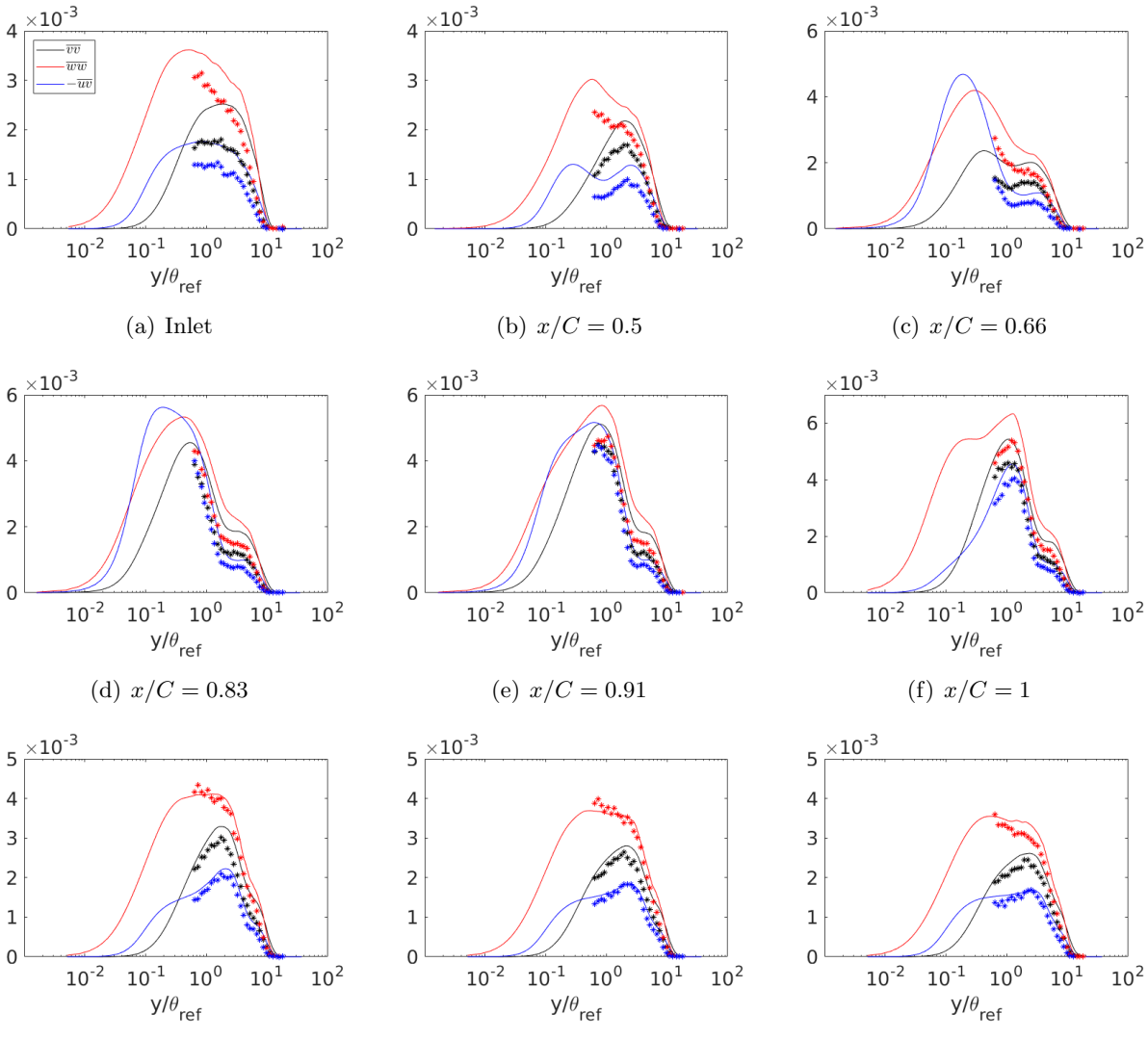
that this peak signals a patch of high turbulent intensity, separating from the surface along the bump. It is enhanced at greater bump heights.

The $\overline{vv}$ and $\overline{ww}$ data are slightly over-predicted by LES. This is similar to what Wu and Squires (1998b) observed. The secondary peak of $\overline{vv}$ develops faster than $\overline{uu}$, emerging at $x/C = 0.66$; but $\overline{ww}$ develops the peak at $x/C = 1$, similarly to $\overline{uu}$. At $x/C = 1.63$, the peak $\overline{vv}$ and $\overline{ww}$ is higher than at the inlet. After the bump, the recovery of both these stresses is not as fast as the mean velocity but the shape of the profile is similar to that upstream.

The $\overline{uv}$ component (fig. 3.9), also, is slightly over-predicted by LES. It develops a second peak $x/C = 0.5$, which is fastest of all the stresses. This peak almost completely disappears by the end of the bump. The magnitude of the $\overline{uv}$ component at $x/C = 1.63$ is similar to that at the inlet, but the profile is very different.

The $\overline{uu}$ peak is an order of magnitude larger than the other diagonal components of the Reynolds stress tensor. The peaks of $\overline{uu}$ at $y_+ = 10$ are at the position of highest production in a zero pressure-gradient flat-plate boundary layer ($y_+ = 11$) as shown in fig 3.6.

All the components tend to develop a second peak over the bump, however these deviations from the flat plate profiles develop at different streamwise locations for the different components. The inner peaks over the bump in the second order moments are not visible in the experimental data because the x-wire measurements did not extend close enough to the wall. These 'knee points' in the profile are indicative of internal layers. These layers are expected because of the steep pressure gradients seen in fig. 3.10 at the start and end of the bump. Vorticity is created or destroyed at the wall (Batchelor, 1967) and this effect diffuses into the flow to create the response layer. Thus, in response to this change in pressure gradient, a new layer is formed within the main boundary layer.



(a) Inlet (b) $x/C = 0.5$ (c) $x/C = 0.66$

(d) $x/C = 0.83$ (e) $x/C = 0.91$ (f) $x/C = 1$

(g) $x/C = 1.33$ (h) $x/C = 1.5$ (i) $x/C = 1.63$

Figure 3.9 Time and span averaged $\overline{vv}$, $\overline{ww}$ and $-\overline{uv}$ at various x locations normalized by U^2_{ref}. —— LES, $*$ experiment.

3.3 Flow Analysis

3.3.1 Base case

The surface static pressure coefficient and non-dimensionalized gradient of pressure are plotted in fig. 3.10. The pressure gradient is adverse as the flow approaches the bump and then becomes favourable until the crest of the bump. It then becomes adverse and drops to being favourable in the recovery region of the flow.

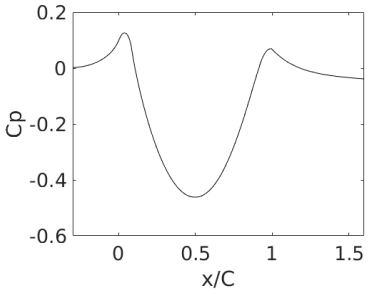
The effect of the sudden change in pressure gradient at the start and end of the bump can be seen on the skin friction coefficient; it too exhibits sudden increase and decrease at $x/C = 0$ and $x/C = 1$. This sudden change in pressure gradient and its effect on the generation of internal layers has been discussed in great detail by Wu and Squires (1998b) and Baskaran et al. (1987) amongst others.

The skin friction coefficient shows a plateau from $x/C = 0.6$ to 0.8 (fig. 3.11). This is not captured by any variation of two equation turbulence models, $k - \omega$ or $k - \epsilon$. The plateau in C_f (observed in this case of $Re_\theta = 2,500$) was attributed by Wu and Squires (1998b) to intermittent flow reversal at the wall, since at $Re_\theta = 1,500$ the streamwise locations of the C_f plateau and the intermittent flow reversal coincide. This, however, is not the case at the current Reynolds number. Fig. 3.10 shows that the intermittent flow reversal at the wall occurs at $x/C = 0.8$ which is after the plateau. Intermittent flow reversal is the percentage of time the flow reverses direction (this was computed by noting the fraction of time the skin friction had a negative sign). The non-dimensional pressure gradient parameters in fig. 3.10 are defined as

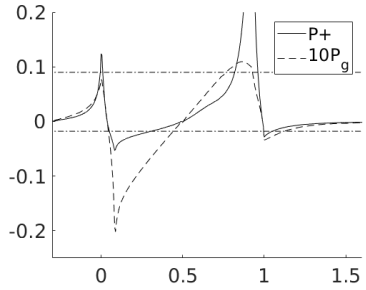
$$P_g = \frac{d\overline{p}_w}{ds}\frac{\theta_{ref}}{U_{ref}^2}; \quad P^+ = \frac{d\overline{p}_w}{ds}\frac{\nu}{\tau_w^{3/2}} \tag{3.5}$$

Fig. 3.10 also marks $P^+ = 0.09$, suggested by Patel (1965) for the onset of separation, and this point marks the beginning of intermittent flow reversal at the wall.

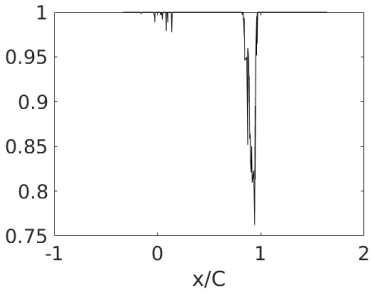
Patel (1965) suggested that $P^+ = -0.018$ marks the onset of relaminarization. The pressure gradient parameter crosses this mark, briefly, just near the front of the bump. In order to verify



(a) C_p

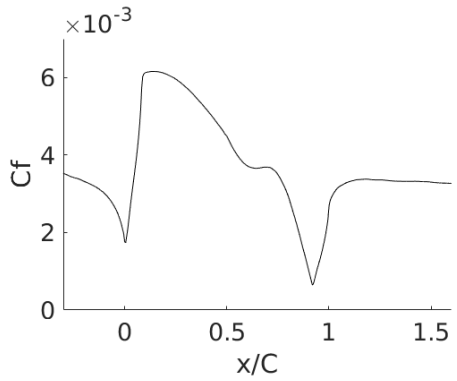


(b) P_g, P_+

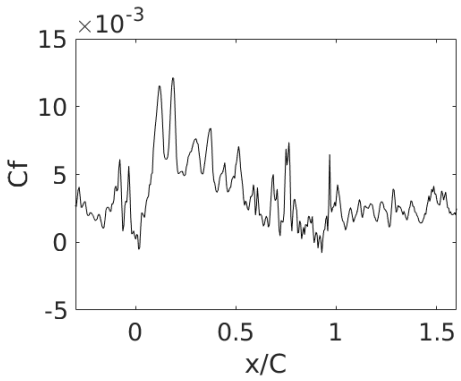


(c) intermittent flow reversal at wall

Figure 3.10 Pressure Coefficient, Pressure gradient parameters and intermittent flow reversal at wall.



(a) C_f



(b) Instantaneous C_f

(c) Instantaneous Skin Friction Coefficient Contours.

Figure 3.11 Skin Friction Coefficient.

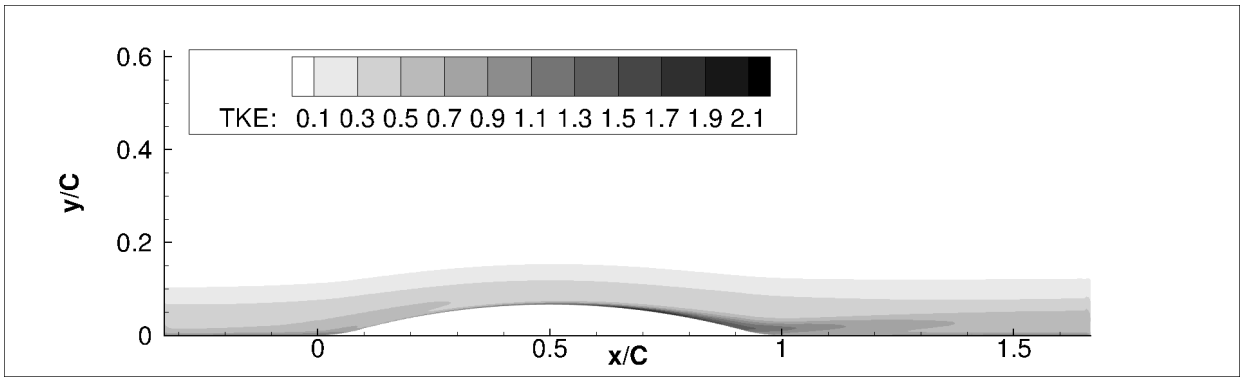


Figure 3.12 TKE contours multiplied by $10^2/U_{ref}^2$.

that the flow does not become laminar over the bump, the instantaneous C_f at the mid span plane is plotted in fig. 3.11. This and the instantaneous skin friction coefficient contours (fig. 3.11) show turbulent fluctuations.

The plateau in the C_f, around $x = 0.7$, is located at a region of high TKE near the wall (fig. 3.12). The high TKE region becomes thicker downstream of the bump. The major contributor to TKE is the $\overline{uu}$ component, and this is shown by fig. 3.13.

The production of $\overline{uu}$ (P_{11}) is an order of magnitude larger than the production of $\overline{vv}$ (P_{22}). The high production starts near the wall and extends into the recovery region, similar to the TKE. This zone of high TKE at the wall, near the plateau of C_f, is caused by a combination of production and turbulent transport: see fig. 3.15. Around the C_f plateau, the near wall turbulent transport is about 50% of the production and is equal to or higher than the inlet production. The reverse is true over the FPG region as the turbulent transport works to decrease the TKE with the local peak value close to the local dissipation of TKE. Production reaches its peak value over the C_f plateau, after which it decreases. Over the APG (adverse pressure gradient) region, the production spreads out. On the contrary, over the FPG region the production region narrows, with almost no production over $y_+ = 50$. The other quantities did not show any remarkable behaviour. The distribution of velocity gradients (fig. 3.14) shows how the velocity gradients extend farther from the surface, on the lee side of the bump. The patch of high TKE is created as the turbulent boundary layer passes through the region of shear, away from the wall.

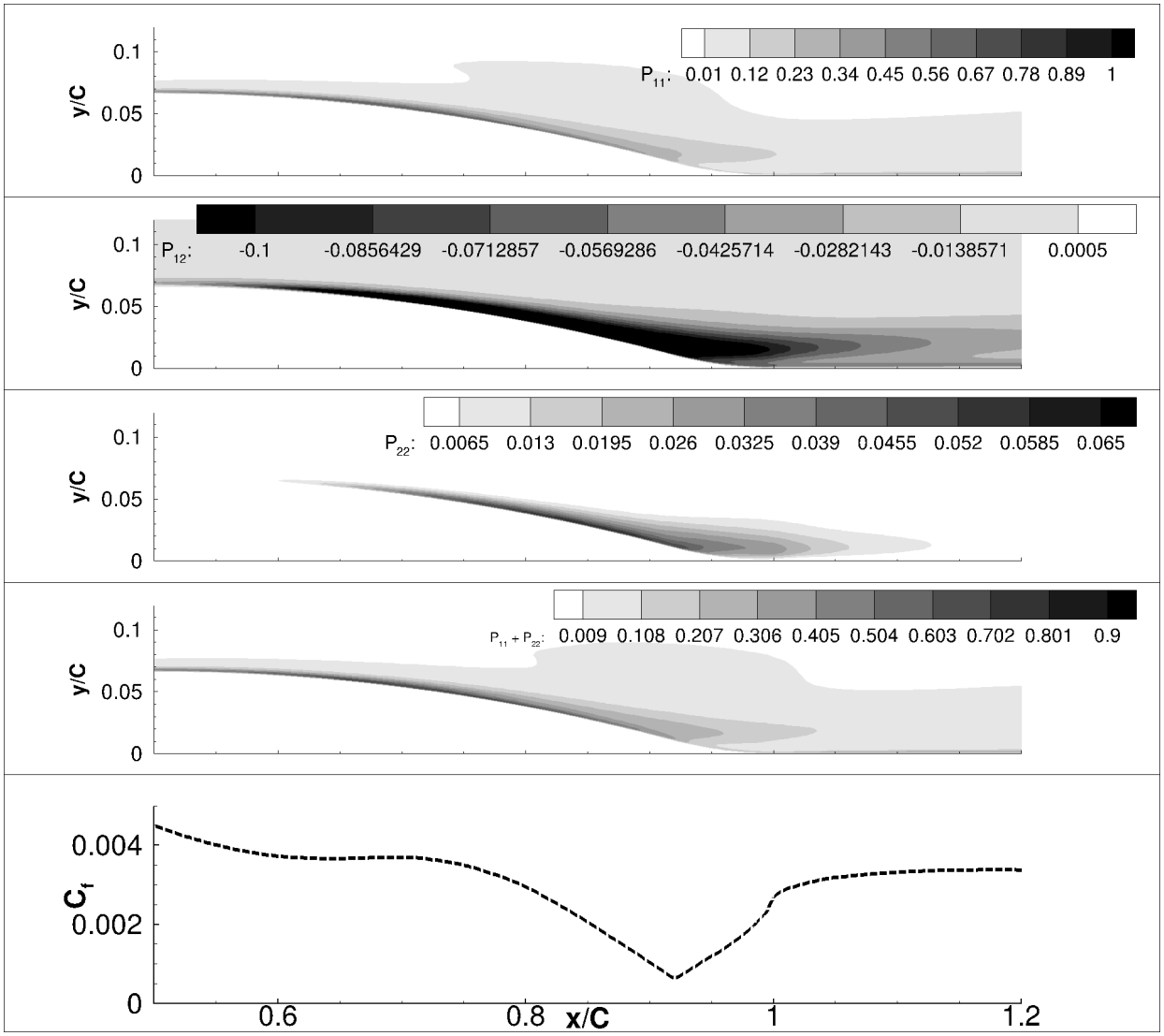


Figure 3.13 Production tensor component contours multiplied by $10^2\theta/U_{ref}^3$. Top to bottom: P_{11}, P_{12}, P_{22}, $P_{11}+P_{22}$. C_f vs. x/C is shown for comparison.

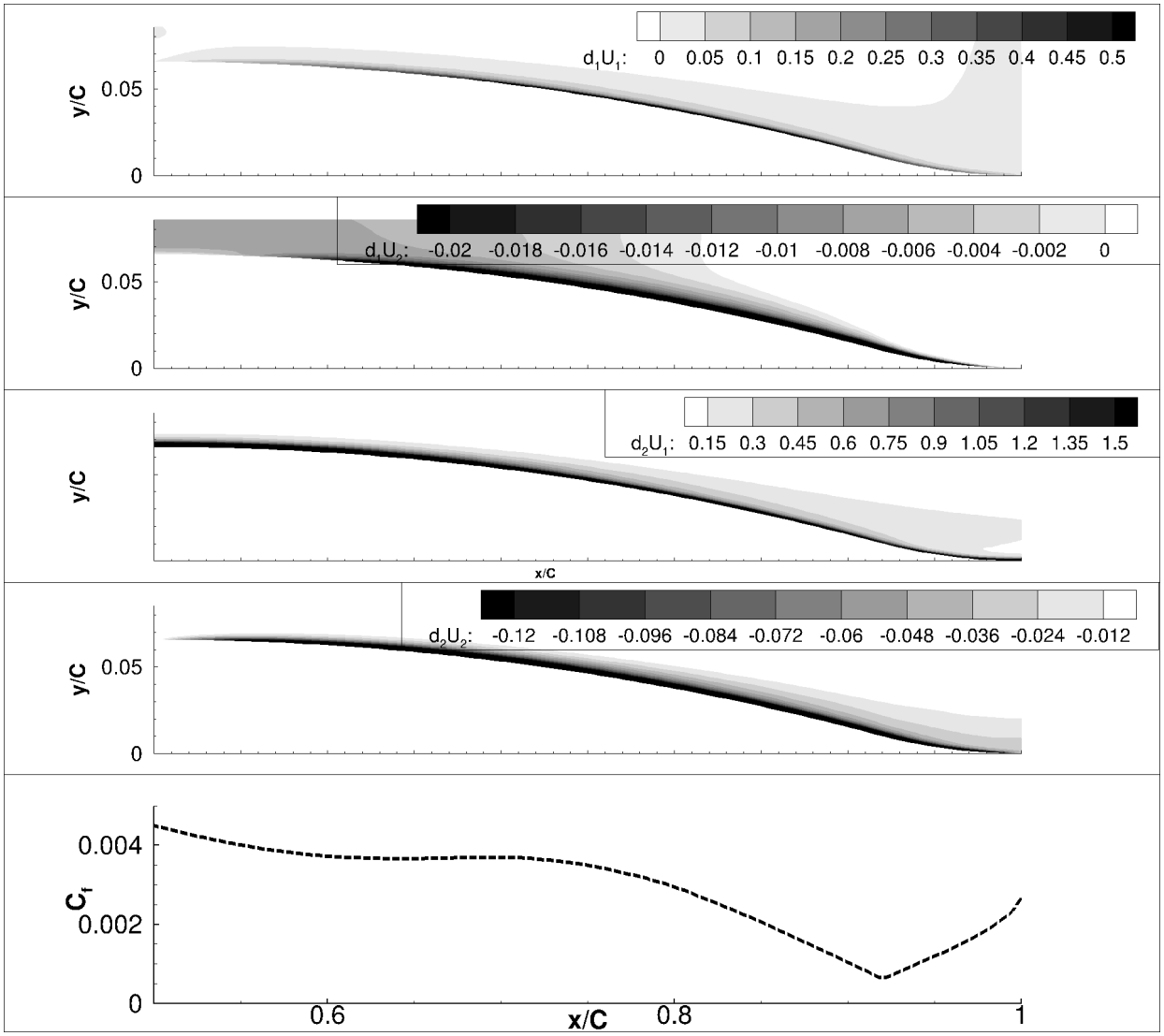
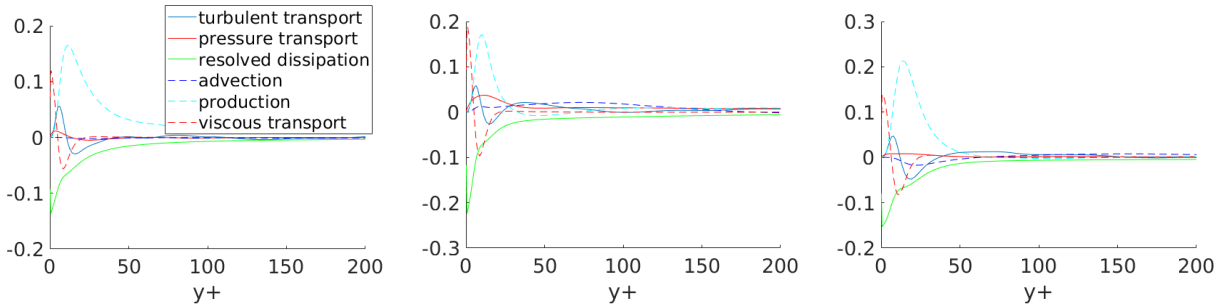
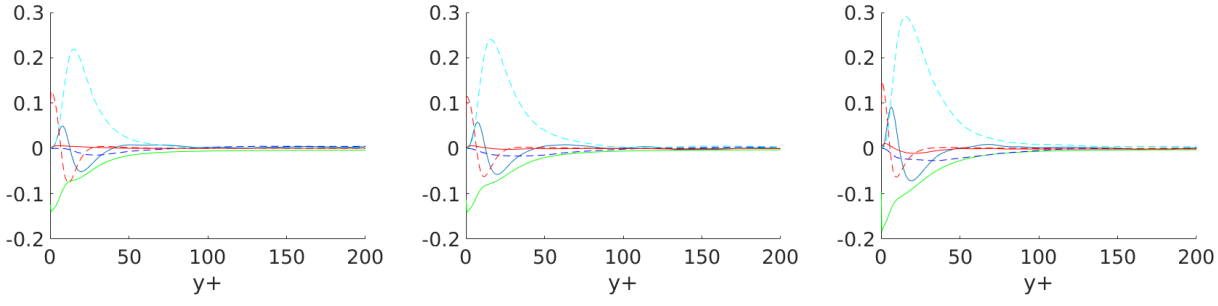


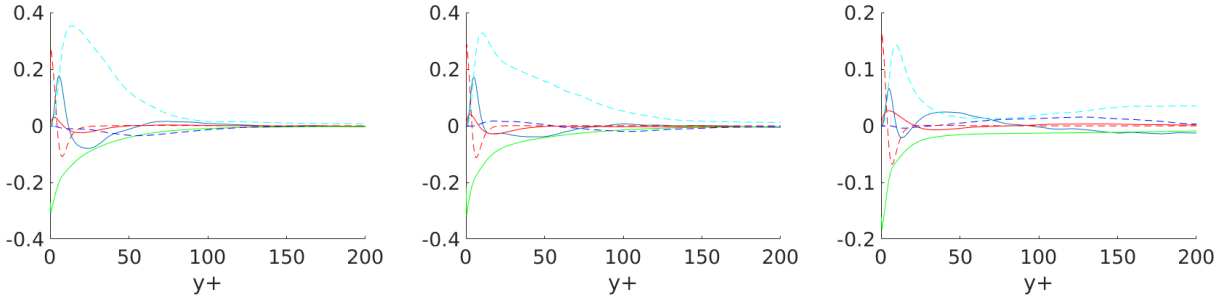
Figure 3.14 Velocity gradient components. multiplied by θ/U_{ref}. Top to bottom: $\partial_1 U_1$, $\partial_1 U_2$, $\partial_2 U_1$, $\partial_2 U_2$. C_f vs. x/C is shown for comparison.



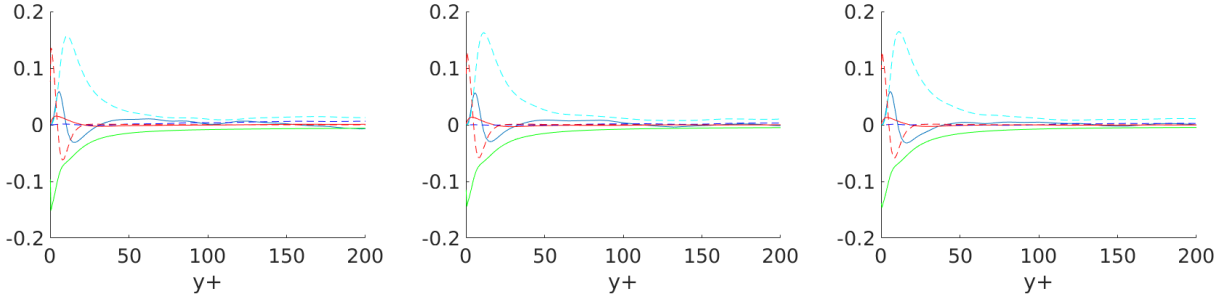
(a) $x/C = -0.26$ (b) $x/C = 0.1$ (c) $x/C = 0.25$



(d) $x/C = 0.35$ (e) $x/C = 0.50$ (f) $x/C = 0.60$



(g) $x/C = 0.70$ (h) $x/C = 0.80$ (i) $x/C = 1.1$



(j) $x/C = 1.33$ (k) $x/C = 1.5$ (l) $x/C = 1.63$

Figure 3.15 TKE budget at various x locations. Terms multiplied by $10^2\theta/U_{ref}^3$.

3.3.2 Bump series

In order to study the progression toward separation, the flow over bumps of increasing height was simulated. The bump geometries are shown in fig. 3.16. The set of bumps is summarized in table 4.3.

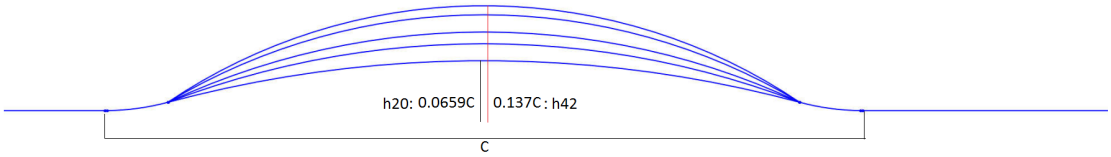


Figure 3.16 Bump geometries.

Table 3.3 Summary of bumps.

Bump	*Height(mm)*	*Description*	*Bump designation*	δ_{99}^{inlet}/R
1	20	Original bump (no separation)	h20	0.069
2	26	flow on verge of separation	h26	0.095
3	31	separated flow	h31	0.113
4	38	separated flow	h38	0.137
5	42	separated flow	h42	0.149

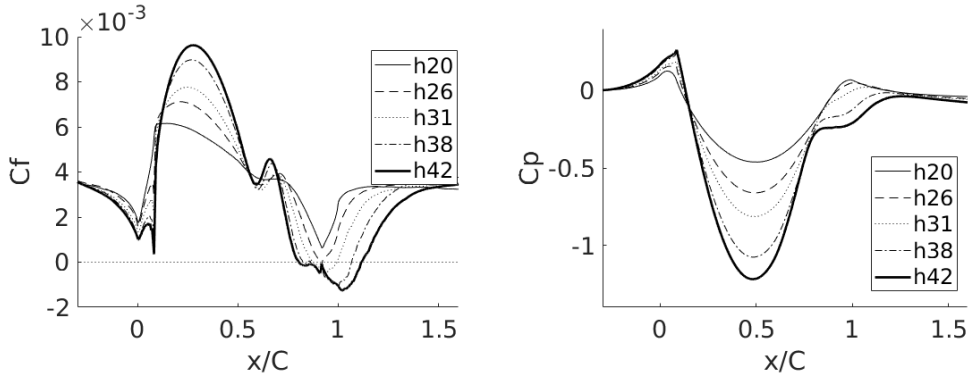
The flow over the original bump, h20, did not separate; h26 is on the verge of separation; and h31-42 develop a small separated region near the end of the bump, as seen in the skin friction coefficient (C_f) plots in fig. 3.17. As the bump height is increased, C_f over the bump also increases, due to greater flow acceleration over the crest, as exemplified by the C_p curves and the pressure gradient curves in fig. 3.18. Note that since P^+ uses τ_w in the denominator the parameter behaves erratically in the separated region of the flow.

P_g becomes progressively more negative as the bump height increases. The region of intermittent flow reversal becomes wider and the percentage time the flow is reversed also increases. The maximum flow reversal increases from 25% of the time for h20 to 100% of the time for h42. Thus

for h42, there exists a region where flow is reversed for the entire time. There is also some reversal where the flow encounters APG before the start of the bump. The C_f plateau near $x/C = 0.7$ becomes a local maximum and a separated region, of negative C_f, develops near $x/C = 0.9$ as the bump height increases.

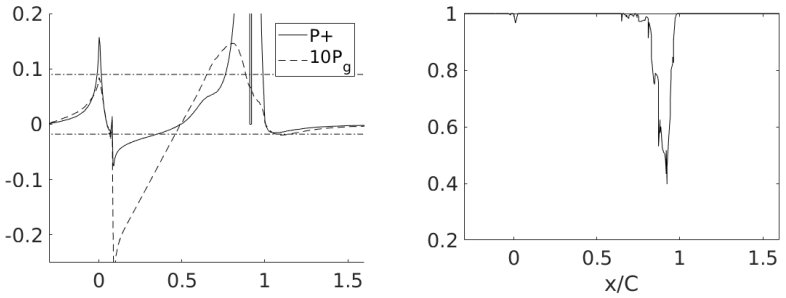
To test sensitivity to incident boundary layer thickness, two additional non-dimensional momentum thicknesses equal to $C/52$ (momentum thickness 2) and $C/101$ (momentum thickness 3) were prescribed at the inlet. The effect of the change in momentum thickness is shown in fig. 3.19. The skin friction coefficient and pressure coefficient profiles remain qualitatively similar; thus, the qualitative behavior is not sensitive to the inlet boundary layer thickness. Note that this sensitivity was tested on mesh 1 of Table 4.1, since the solutions on meshes 3 and 1 were nearly identical for the original/baseline bump (h20).

The instantaneous skin friction coefficient in a z-plane, and its surface contours are plotted in fig. 3.20. The contours show streaky structures and the line plots show turbulent structures, suggesting that the flow did not relaminarize even over the highest bump. It is to be noted that even though the magnitude of $\overline{uu}$ was an order of magnitude greater than $\overline{vv}$ and $\overline{ww}$ (as will be seen later), the velocity field did not become a single scale velocity field associated with a laminar flow.

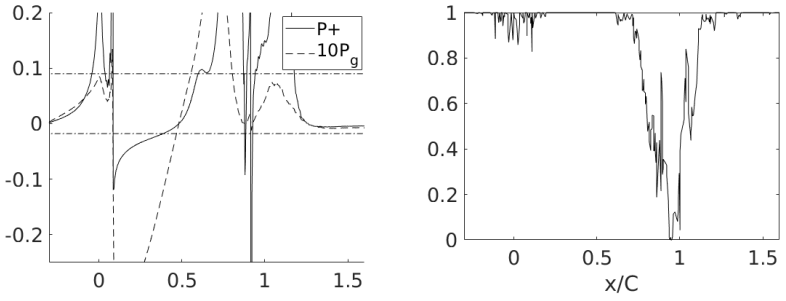


(a) C_f (b) C_p

Figure 3.17 Comparison of skin friction coefficient and pressure for different bump crest heights.

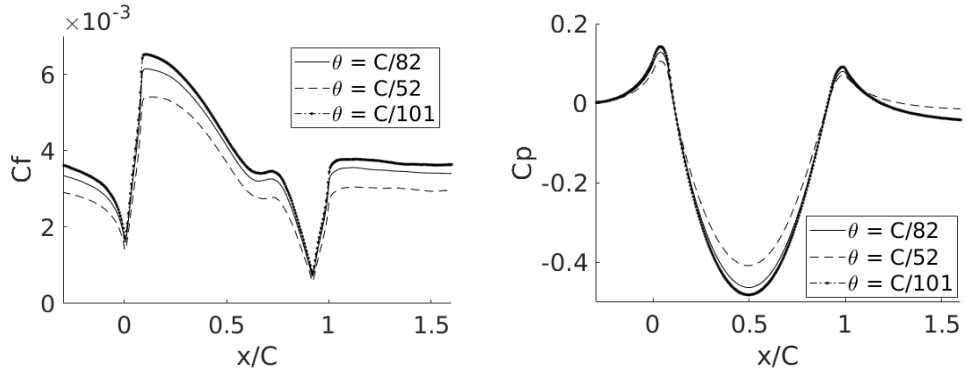


(a) h26



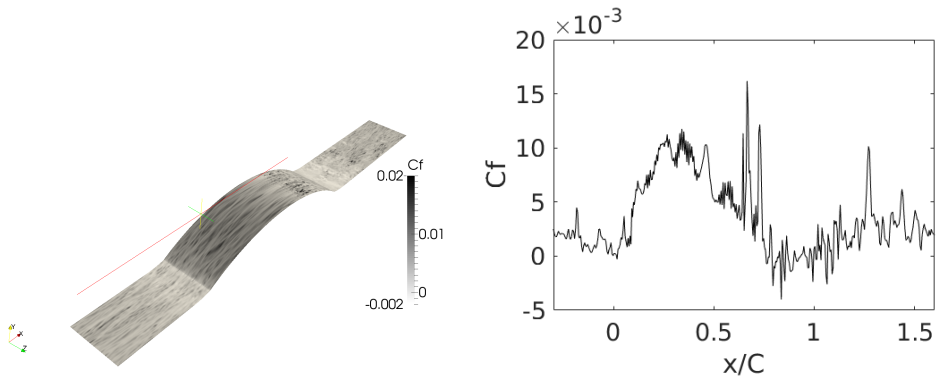
(b) h42

Figure 3.18 Pressure gradient parameters and intermittent flow reversal at wall for bumps with different crest heights.



(a) C_f (b) C_p

Figure 3.19 Comparison of skin friction coefficient and pressure for different inlet momentum thicknesses (h20).



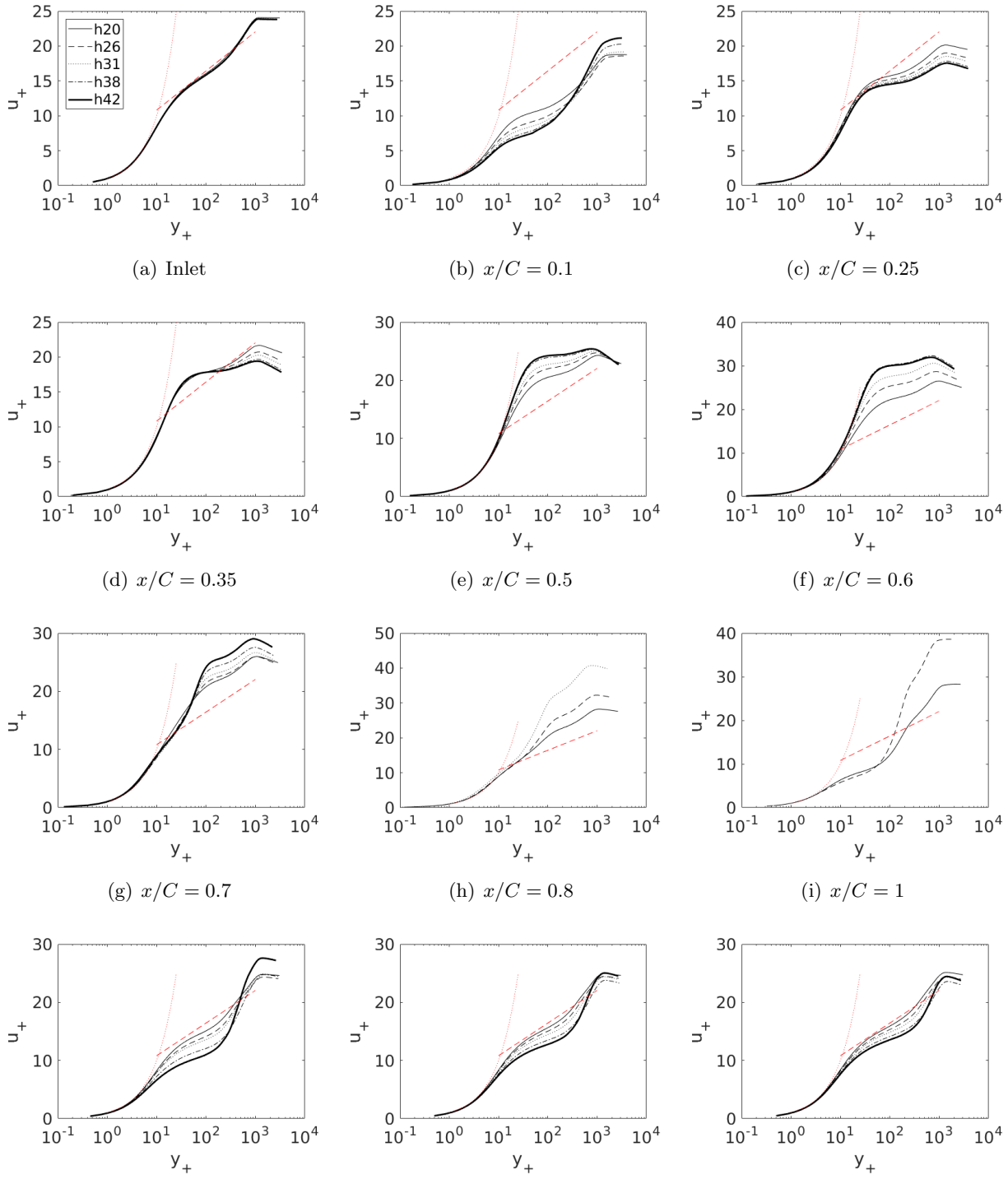
(a) Instantaneous C_f contours (b) Instantaneous C_f

Figure 3.20 Instantaneous Skin Friction Coefficient for h42.

Fig. 3.21 provides the streamwise mean velocity profiles in viscous units. As in the base simulation, major departure from the log-law can be seen over the bump. From fig. 3.18 it is seen that the flow changes from APG to FPG around $x/C = 0.08$ and an immediate deviation from log law is observed. Profiles after $x/C = 0.1$ show similarity to reverse transitional profiles (Page et al., 1952; Patel and Head, 1968); i. e., we see an increase in the thickness of the viscous sublayer that becomes greater as the bump height increases. However, as noted earlier, the flow is never laminar. It is interesting to note that this thick viscous sublayer continues, even after the pressure gradient changes signs (FPG turns to APG) at $x/C = 0.5$. The present simulations parallel behaviours seen over flat and curved surfaces (Tsuji and Morikawa, 1976): in the experiments of Tsuji and Morikawa (1976), APG did not result in deviation from the log law, but FPG did.

The profiles progressively deviate from h20 and major deviations can be seen after $y_+ = 10$. At $x/C = 0.8$ and $x/C = 1$ profiles have not been plotted for cases in which the flow separated. After flow separation all the profiles tend towards h20 by $x/C = 1.33$. But, by the end of the domain, the flow hasn't fully recovered.

Figs. 3.22, 3.23, 3.24 and 3.25 show components of the Reynolds stress tensor for the complete set of bumps. As before, $\overline{uu}$ is the largest diagonal component of the stress tensor; i.e. it is the major contributor to the TKE, with its highest value being an order of magnitude larger than those of $\overline{vv}$ and $\overline{ww}$.



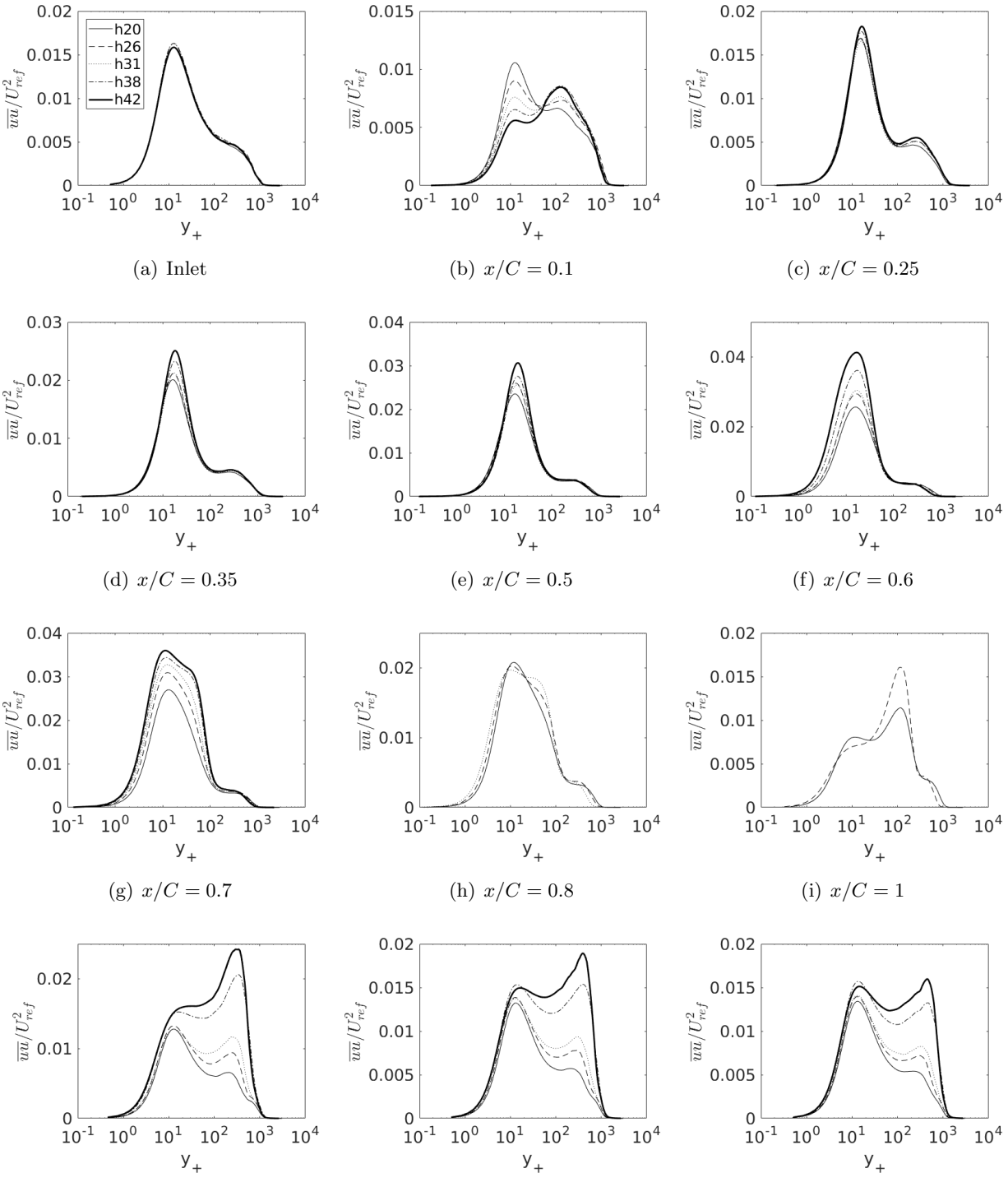
(a) Inlet (b) $x/C = 0.1$ (c) $x/C = 0.25$

(d) $x/C = 0.35$ (e) $x/C = 0.5$ (f) $x/C = 0.6$

(g) $x/C = 0.7$ (h) $x/C = 0.8$ (i) $x/C = 1$

(j) $x/C = 1.33$ (k) $x/C = 1.5$ (l) $x/C = 1.63$

Figure 3.21 Time and span averaged x-velocity at various x locations. —— h20, - - - h26, ····· 31, — · — 38, ▬▬ 42.



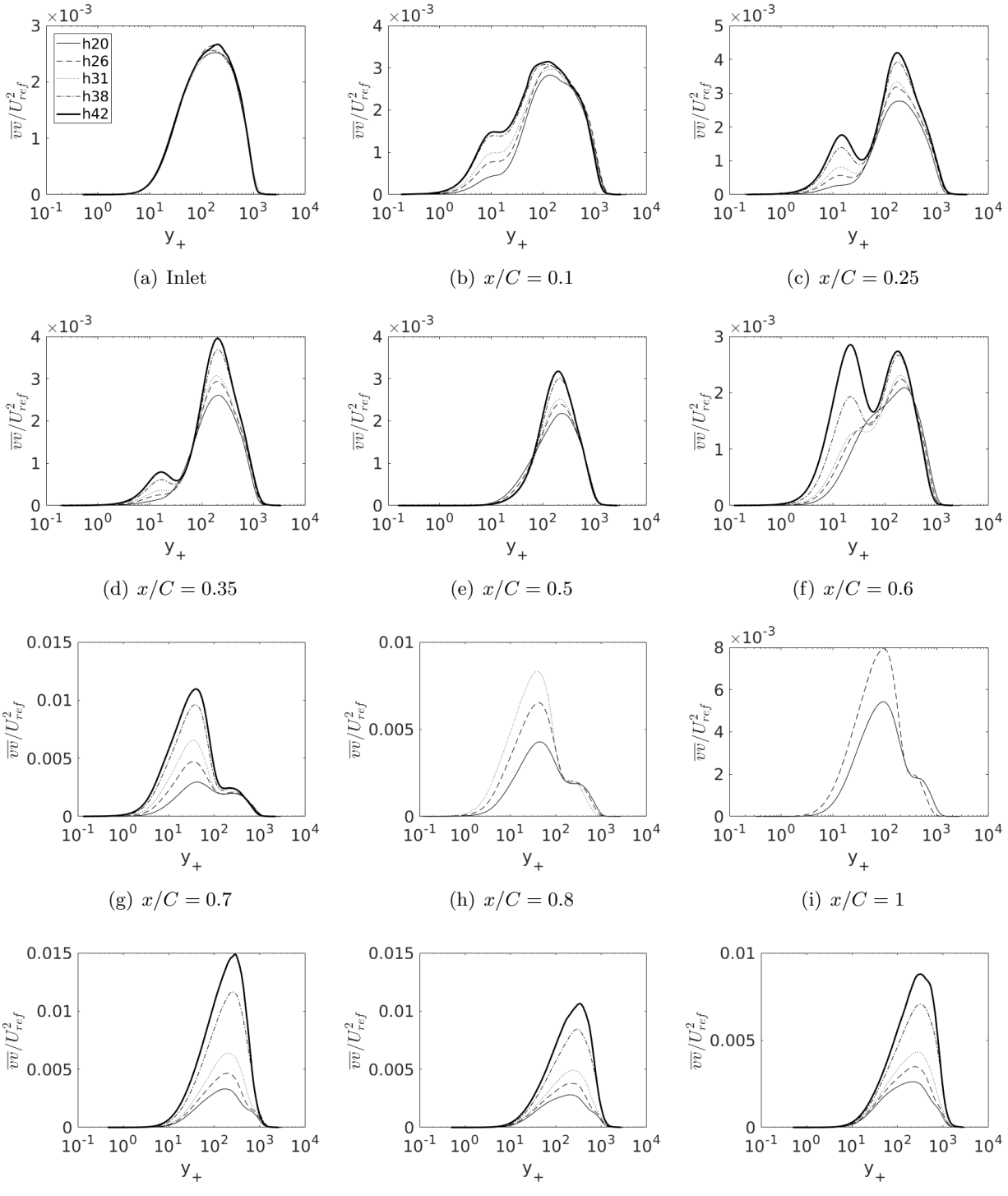
(a) Inlet (b) $x/C = 0.1$ (c) $x/C = 0.25$

(d) $x/C = 0.35$ (e) $x/C = 0.5$ (f) $x/C = 0.6$

(g) $x/C = 0.7$ (h) $x/C = 0.8$ (i) $x/C = 1$

(j) $x/C = 1.33$ (k) $x/C = 1.5$ (l) $x/C = 1.63$

Figure 3.22 Time and span averaged $\overline{uu}$ at various x locations. —— h20, - - - h26, ····· 31, — · — 38, ▬▬ 42.



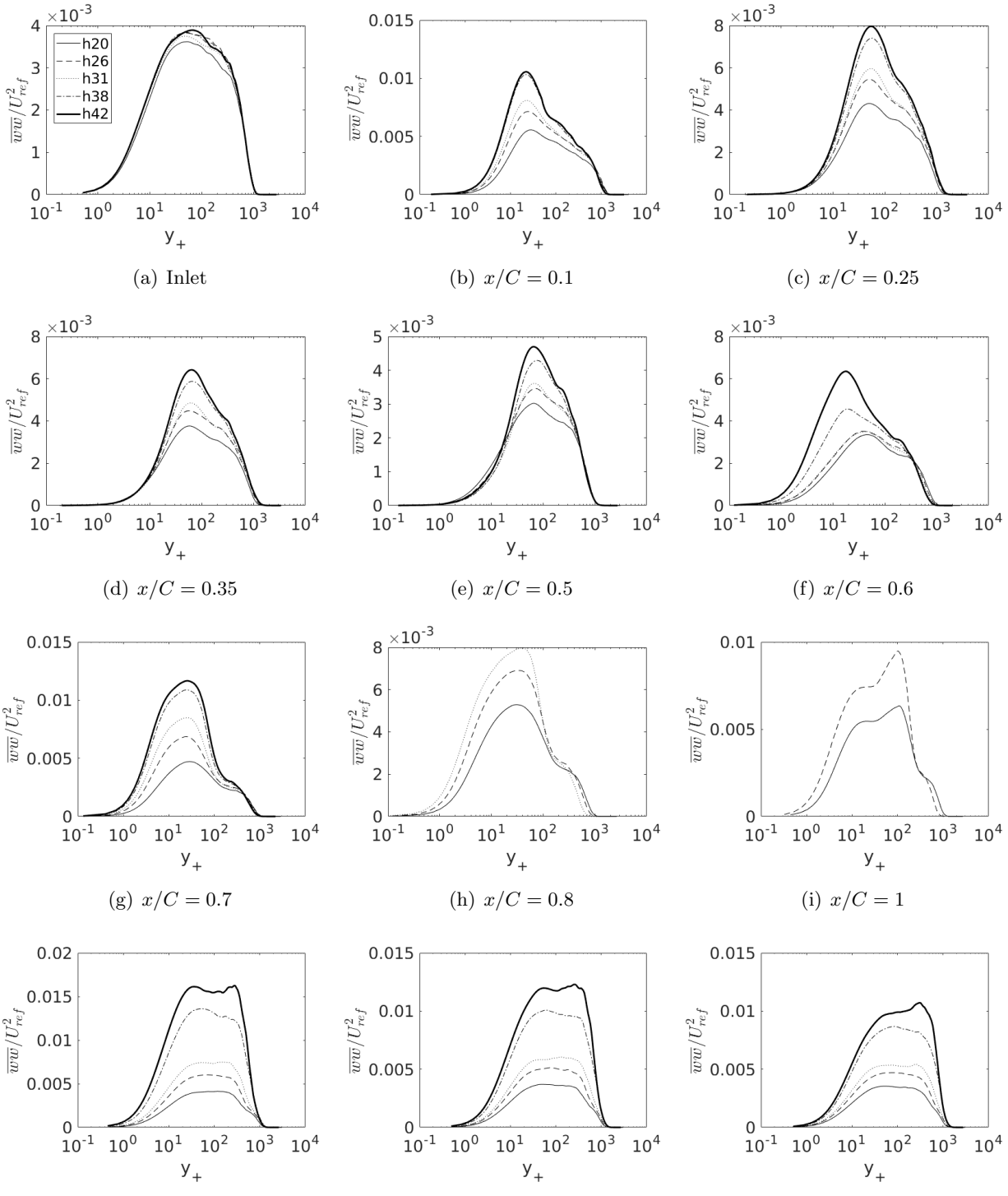
(a) Inlet (b) $x/C = 0.1$ (c) $x/C = 0.25$

(d) $x/C = 0.35$ (e) $x/C = 0.5$ (f) $x/C = 0.6$

(g) $x/C = 0.7$ (h) $x/C = 0.8$ (i) $x/C = 1$

(j) $x/C = 1.33$ (k) $x/C = 1.5$ (l) $x/C = 1.63$

Figure 3.23 Time and span averaged $\overline{vv}$ at various x locations. —— h20, - - - h26, ····· 31, — · — 38, ▬▬ 42.



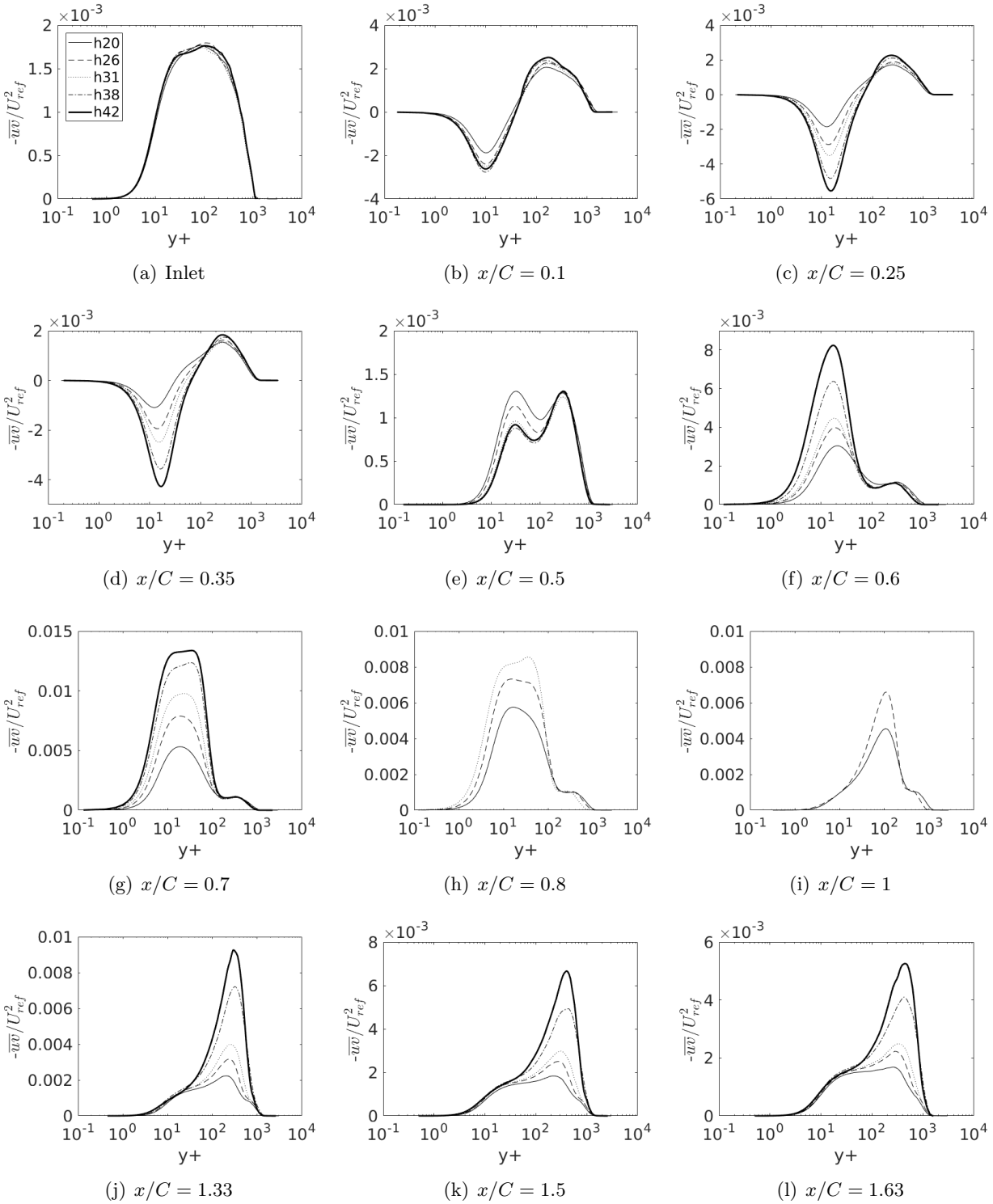
(a) Inlet (b) $x/C = 0.1$ (c) $x/C = 0.25$

(d) $x/C = 0.35$ (e) $x/C = 0.5$ (f) $x/C = 0.6$

(g) $x/C = 0.7$ (h) $x/C = 0.8$ (i) $x/C = 1$

(j) $x/C = 1.33$ (k) $x/C = 1.5$ (l) $x/C = 1.63$

Figure 3.24 Time and span averaged $\overline{ww}$ at various x locations. —— h20, - - - h26, ····· 31, — · — 38, ▬▬ 42.



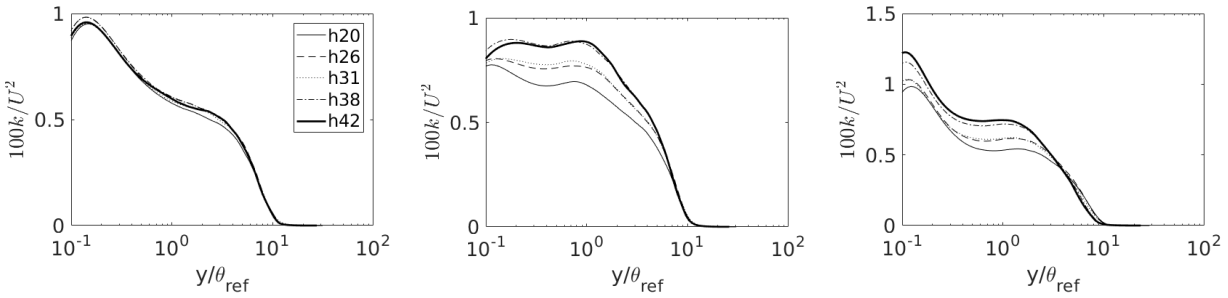
(a) Inlet (b) $x/C = 0.1$ (c) $x/C = 0.25$

(d) $x/C = 0.35$ (e) $x/C = 0.5$ (f) $x/C = 0.6$

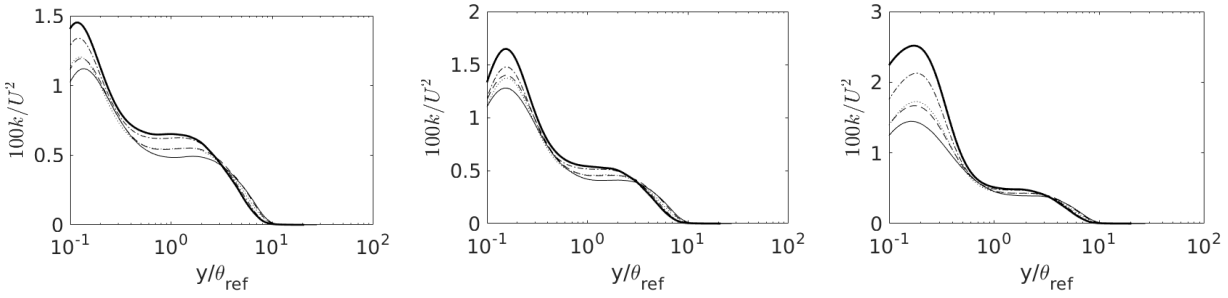
(g) $x/C = 0.7$ (h) $x/C = 0.8$ (i) $x/C = 1$

(j) $x/C = 1.33$ (k) $x/C = 1.5$ (l) $x/C = 1.63$

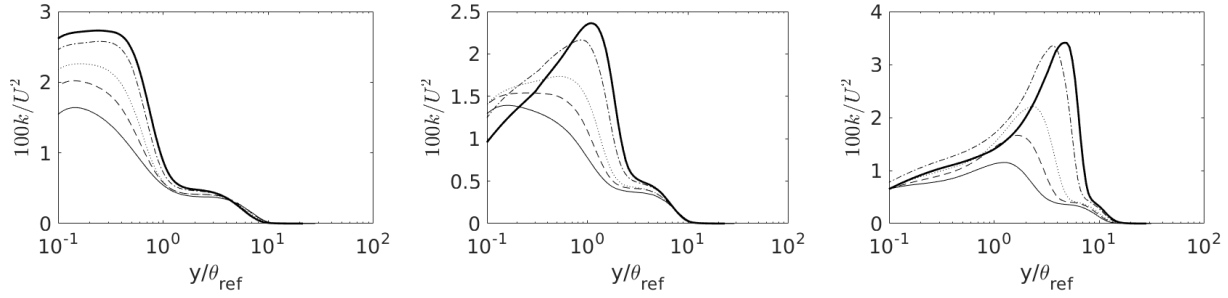
Figure 3.25 Time and span averaged -$\overline{uv}$ at various x locations. —— h20, - - - h26, ····· 31, — · — 38, ▬▬ 42.



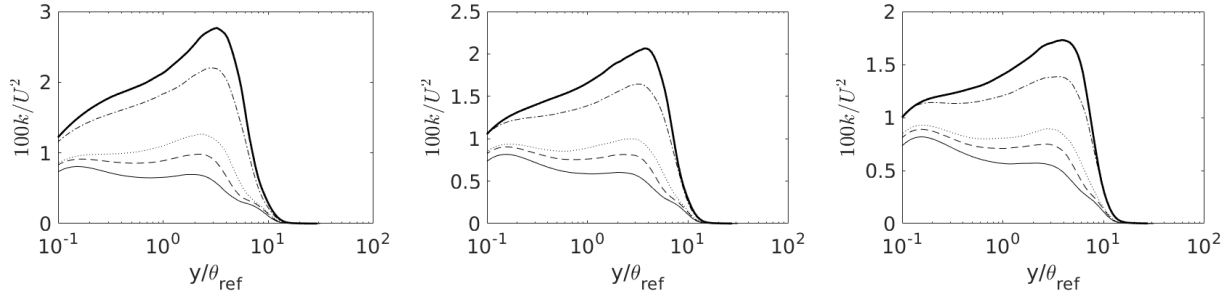
(a) Inlet (b) $x/C = 0.1$ (c) $x/C = 0.25$



(d) $x/C = 0.35$ (e) $x/C = 0.5$ (f) $x/C = 0.6$

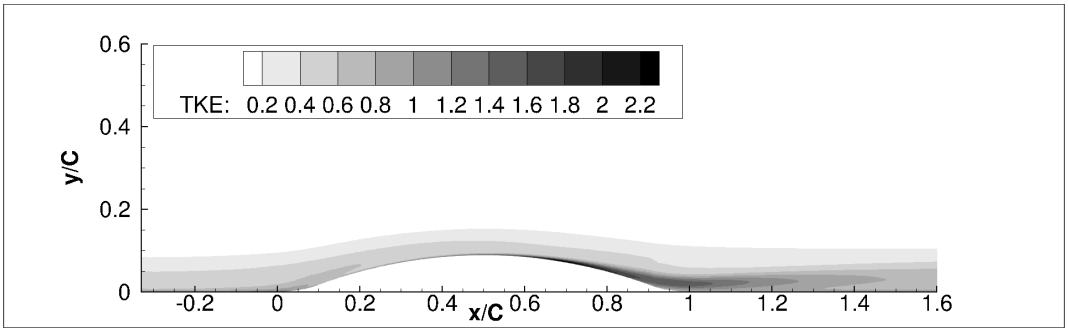


(g) $x/C = 0.7$ (h) $x/C = 0.8$ (i) $x/C = 1$

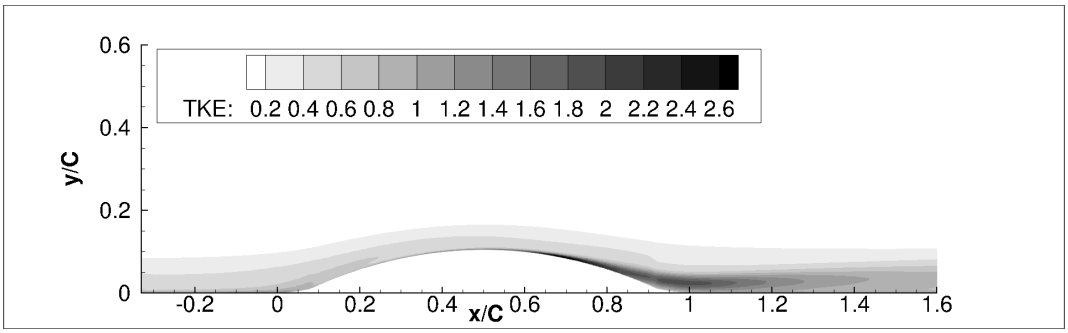


(j) $x/C = 1.33$ (k) $x/C = 1.5$ (l) $x/C = 1.63$

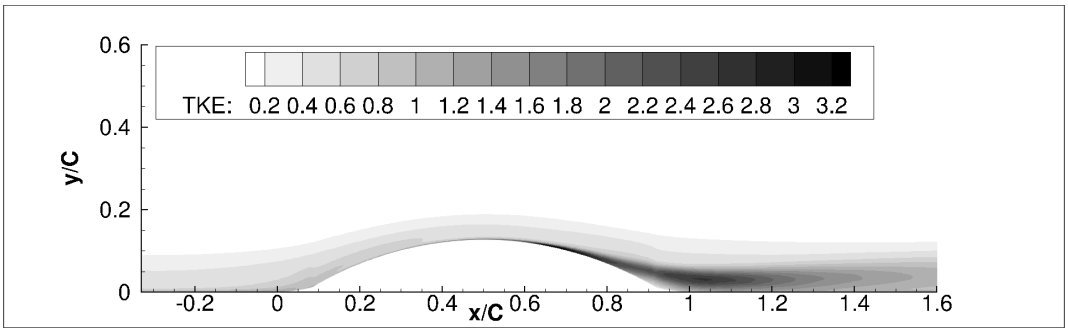
Figure 3.26 $10^2 k/U^2$ at various x locations. —— h20, - - - h26, ····· 31, — · — 38, ▬▬ 42.



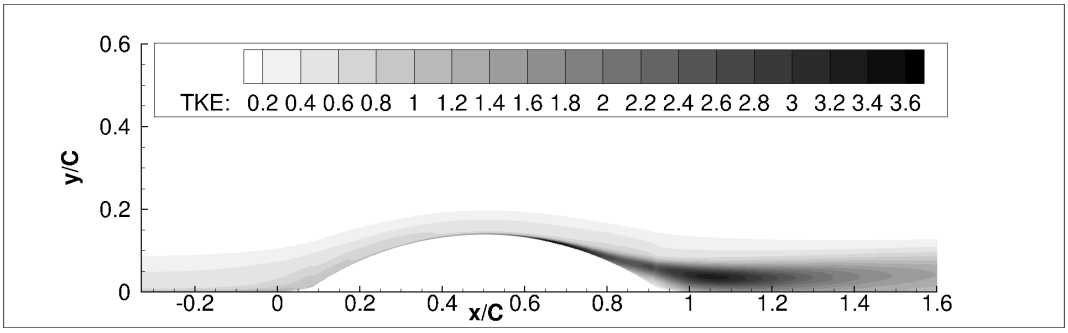
(a) h26



(b) h31



(c) h38



(d) h42

Figure 3.27 TKE contours multiplied by $10^2/U_{ref}^2$.

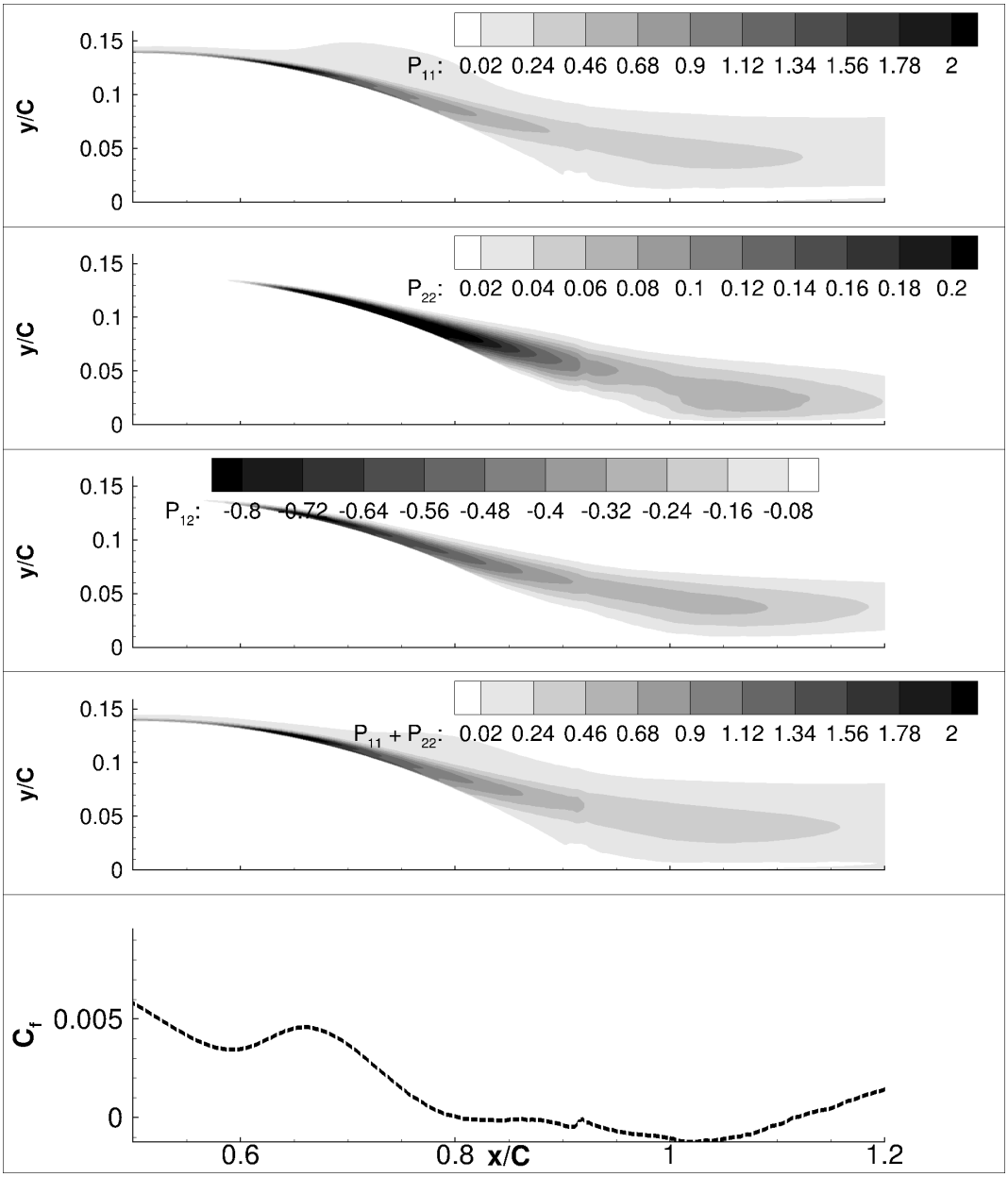
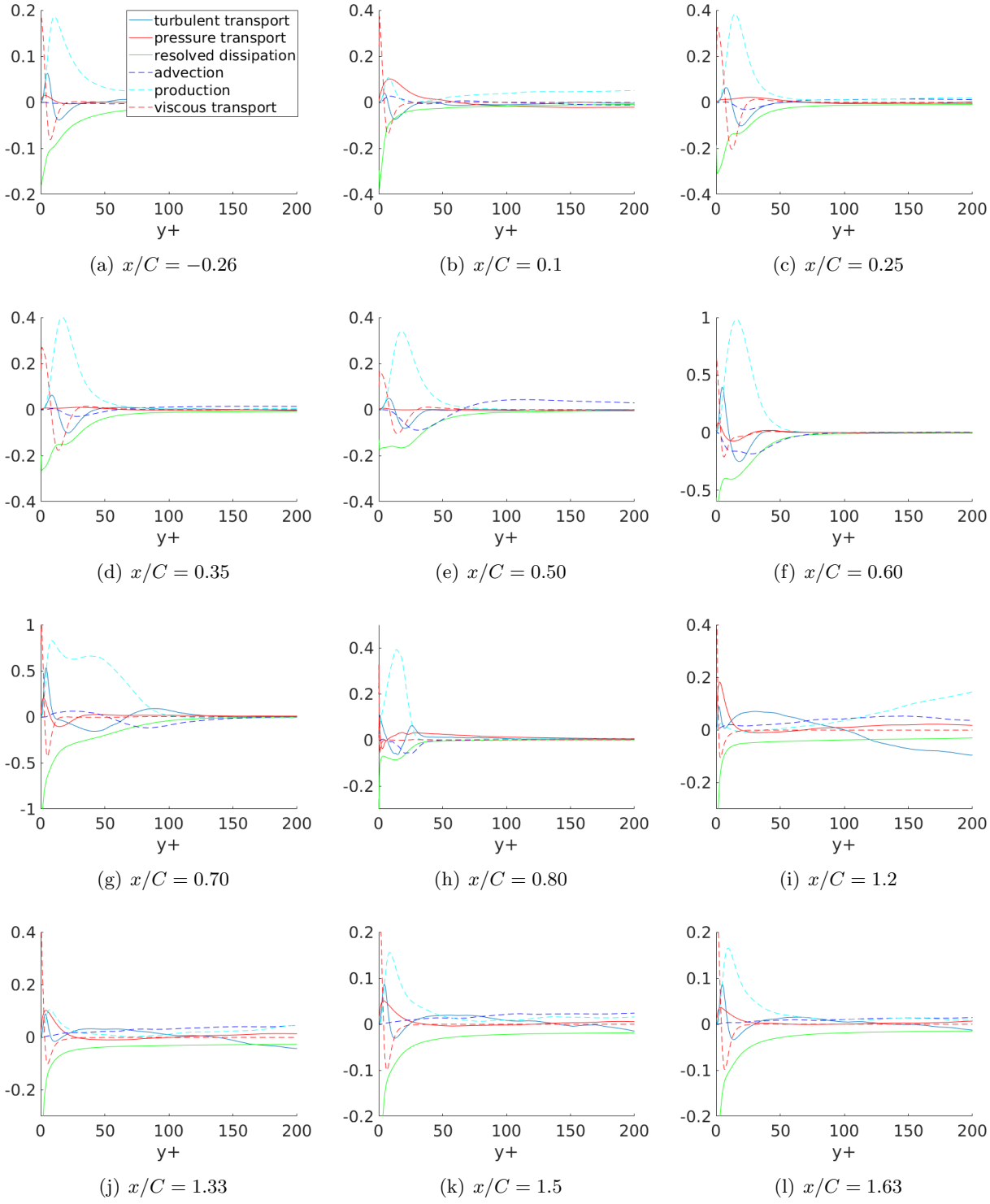


Figure 3.28 Production tensor component contours for h42 multiplied by $10^2\theta/U_{ref}^3$. Top to bottom: P_{11}, P_{22}, P_{12}, $P_{11}+P_{22}$. C_f vs. x/C is shown for comparison.



(a) $x/C = -0.26$ (b) $x/C = 0.1$ (c) $x/C = 0.25$

(d) $x/C = 0.35$ (e) $x/C = 0.50$ (f) $x/C = 0.60$

(g) $x/C = 0.70$ (h) $x/C = 0.80$ (i) $x/C = 1.2$

(j) $x/C = 1.33$ (k) $x/C = 1.5$ (l) $x/C = 1.63$

Figure 3.29 TKE budget for h42 at various x locations. Terms multiplied by $10^2\theta/U_{ref}^3$.

The $\overline{uu}$ profiles show that increasing the crest height decreases the first peak, at around $y_+ = 10$, and increases the second peak in the FPG, at $x/C = 0.1$. Over the crest and in the APG, the inner peak form a plateau. The peaks re-emerge over the recovery region, and the outer peaks are much higher for bumps with higher crest height. For all the bumps, the inner peak for $\overline{uu}$ recovers by $x/C = 1.5$.

The inner plateau of $\overline{vv}$, around $y_+ = 10$, increases with increase in bump height, over FPG. The profiles for various bump heights remain distinct by the end of the domain (66%C after the bump).

The profiles of $\overline{ww}$ are broader than the other components. They are fairly flat from $y_+ = 10$ to $y_+ = 1,000$. Over the FPG they tilt downward from a near-wall maximum. Over the APG, they fill back in, flattening by the end of the bump. However, the profiles for the various bump heights have become quite distinct by that point. The profiles for the higher bumps do not recover by the end of the domain with the plateau being over 140% higher for h42 than at the inlet.

In fig. 3.25, $-\overline{uv}$ develops a negative lobe near the wall under FPG. Hence, this is a region of turbulence suppression. This lobe has returned to positive by $x/C = 0.5$. The level of $-\overline{uv}$ progressively increases as the crest height increases. By $x/C = 0.688$, h42 has a peak value over two times greater than h20. A similar trend can be seen downstream of the bump and the highest bump is the furthest off from the inlet profile. Thus, as seen before, the $\overline{uv}$ component takes a longer distance to recover to the flat plate solution than the mean streamwise velocity profiles.

Fig. 3.26 shows the TKE for all the bumps at various x-locations. The "knee points" can be clearly seen and are formed as a result of surface curvature discontinuity. A knee point is present around $y/\theta_{ref} = 0.3$ at $x/C = 0.1$ and moves away from the wall as the flow progresses downstream, indicating the growth of the internal layer (Webster et al., 1996) as discussed earlier. The flow encounters another curvature discontinuity around $x/C = 0.92$ and another knee point can be observed by $x/C = 1.33$. By this location the knee point from the previous discontinuity is around $y/\theta_{ref} = 7$ and has almost disappeared.

An overview of the TKE field is provided by contour plots in fig. 3.27. A high TKE band originates near the wall, on the lee side. It originates close to the position where the C_f plateau is observed in fig. 3.17 and extends downstream over the recovery region. The band extends further downstream for higher bumps and the magnitude also increases. Note that this is not a region of separated flow. While that behavior might seem extraordinary, it was seen in RANS computations with eddy viscosity. The primary discrepancy was that it did not originate so close to the wall.

The production of $\overline{uu}$ for h42 is an order of magnitude larger than $\overline{vv}$, as shown in fig. 3.28. All the production components start near the wall and extend into the recovery region. It can be seen that the streamwise position of this increase almost perfectly coincides with the increase in C_f which in turn indicates an increase in shear as noted for the base case. The TKE budgets for flow over h42 (fig. 3.29) shows that in the near wall ($y_+ < 20$) APG region, the turbulent transport can be more than 50% of the production. Thus, as with the base case, turbulent transport is a major contributor to near wall TKE. Again, the FPG narrows down the production and increases the negative turbulent transport and negative advection decreasing the TKE. Further away from the wall in the APG region, production remains high almost to $y_+ = 100$, forming a plateau that starts a $y_+ = 10$. In zero and FPG regions, production has a sharp peak at $y_+ \approx 10$. At $x/C = 1.2$ the production peaks far away from the wall around $y_+ = 200$. After that in the recovery region, it collapses rapidly back the zero pressure gradient form.

3.4 Data Extraction by Inversion

A juggernaut to using data to improve RANS models is the ambiguous connection between model variables and measurable quantities. Obvious examples are the eddy viscosity and ω, of the $k-\omega$ model. These are defined by the particular closure model. But, even k, as used in the $k-\omega$ and $k-\epsilon$ models, does not correspond to turbulent kinetic energy — for example, it equals $3.3u_*^2$ in the log-layer, while the actual TKE is higher and depends on Reynolds number (Durbin and Reif, 2011). Duraisamy et al. (2015); Parish and Duraisamy (2016a) proposed an innovative method to

extract data on model variables by solving an optimization problem. The method is summarized below; Singh et al. (2017) also contains an application of the method to the present data.

The optimization is accomplished after introducing a coefficient into the $k-\omega$ model. (Although the coefficient can be inserted in various places, the results were quite similar.) The production term of the ω equation is multiplied by a correction coefficient β in

$$\frac{D\omega}{Dt} = \beta(\boldsymbol{x})\, c_{\omega 1}|S|^2 - c_{\omega 2}\omega^2 + \nabla\cdot(\nu + \sigma_\omega k/\omega)\nabla\omega + (\sigma_d/\omega)\nabla k\cdot\nabla\omega$$

β is initialized to the baseline value of unity. Then values of β over the grid are found by minimizing a cost function, J. Two cost functions were examined: the surface skin friction, and the x velocity profiles; either

$$J = \sum (C_f^{RANS} - C_f^{LES})^2 A_{face}, \quad \text{or} \quad J = \sum (U^{RANS} - U^{LES})^2 \frac{V_{cell}}{d_w^2} \tag{3.6}$$

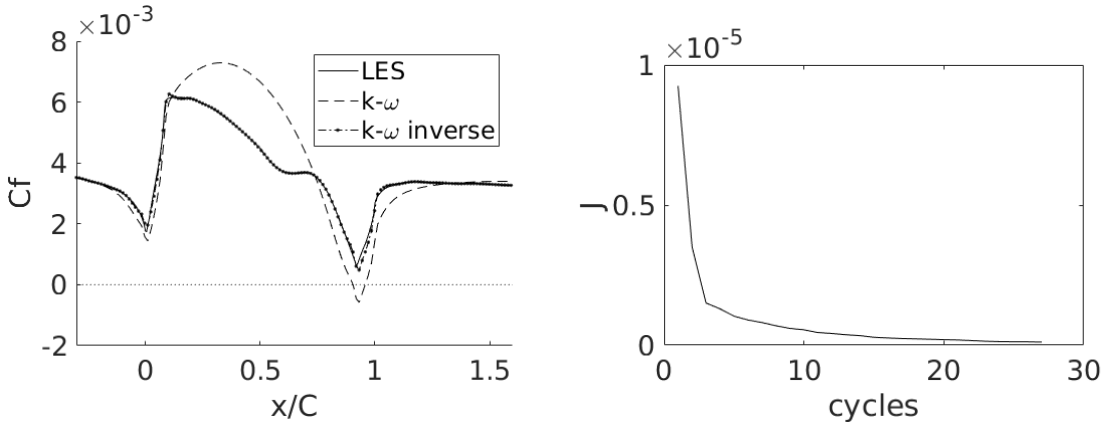
where d_w is wall distance.

The optimization problem is $\min_\beta J$. This requires solving the full RANS equations, with the eddy viscosity equal to k/ω, at each iteration of the minimization. A gradient descent algorithm was used to reach the optimum solution. The gradient of the cost ($\partial J/\partial\beta$), which represents the sensitivity of the objective function to β, was calculated using the adjoint method. In formal terms, the adjoint vector ϕ and the gradient of the cost function with respect to β are found from the system

$$\frac{\partial R}{\partial \boldsymbol{U}}^T \phi = -\frac{\partial J}{\partial \boldsymbol{U}}^T; \quad \frac{dJ}{d\beta} = \phi^T \frac{\partial R}{\partial \beta} \tag{3.7}$$

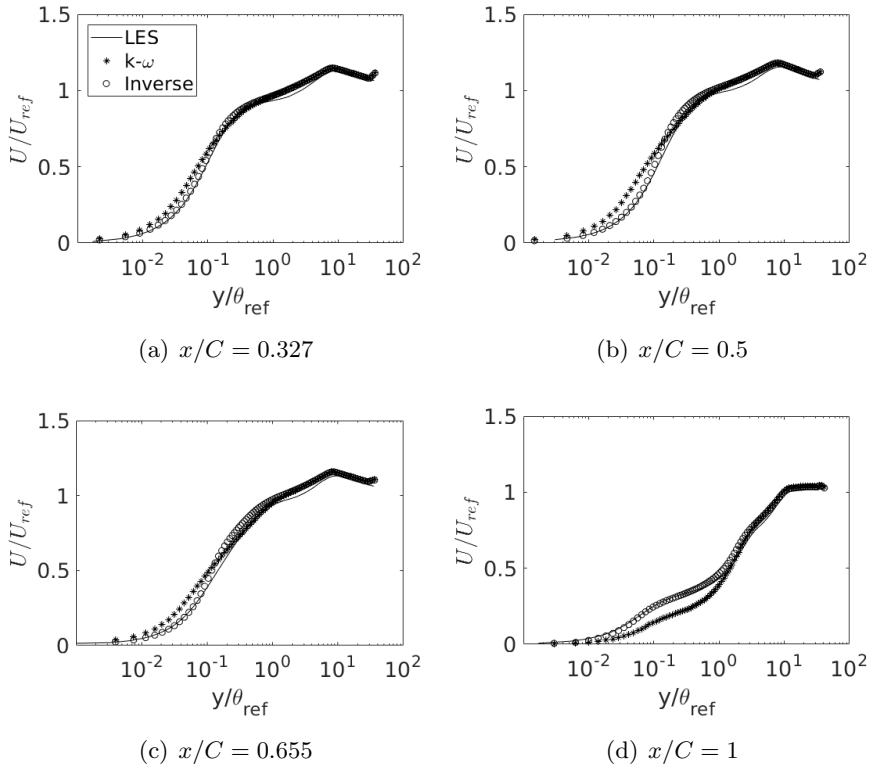
where R = 0 are the governing equations and $\boldsymbol{U}$ is a vector all the variables in the RANS equations (see Duraisamy et al. (2015) for a more complete discussion).

Fig. 3.30 shows that once the optimum β was reached (in about 20 cycles), the C_f matched almost exactly with the LES data. Fig. 3.31 shows that, even with C_f as the cost function, the velocity profiles match well with the LES data; thus, although the cost function is a wall quantity its effect on the solution is not local. Fig. 3.32 shows that corrections are needed, mostly near the wall.



(a) C_f (b) Objective Function

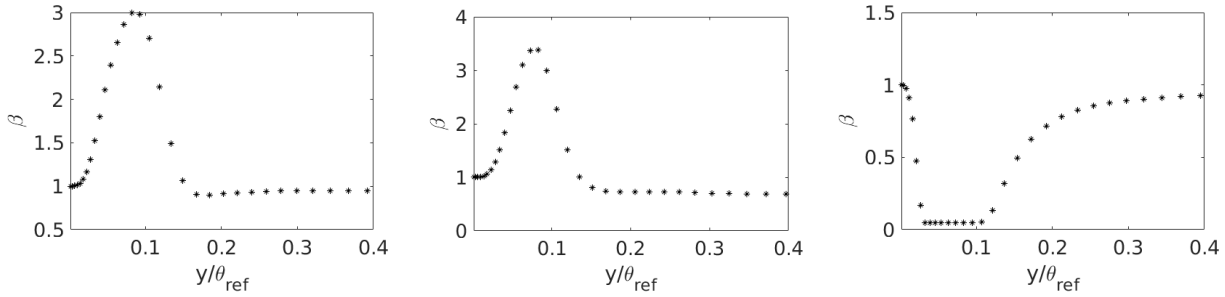
Figure 3.30 C_f of inverse solution compared to LES and RANS solutions and the objective function for the optimization (h20).



(a) $x/C = 0.327$ (b) $x/C = 0.5$

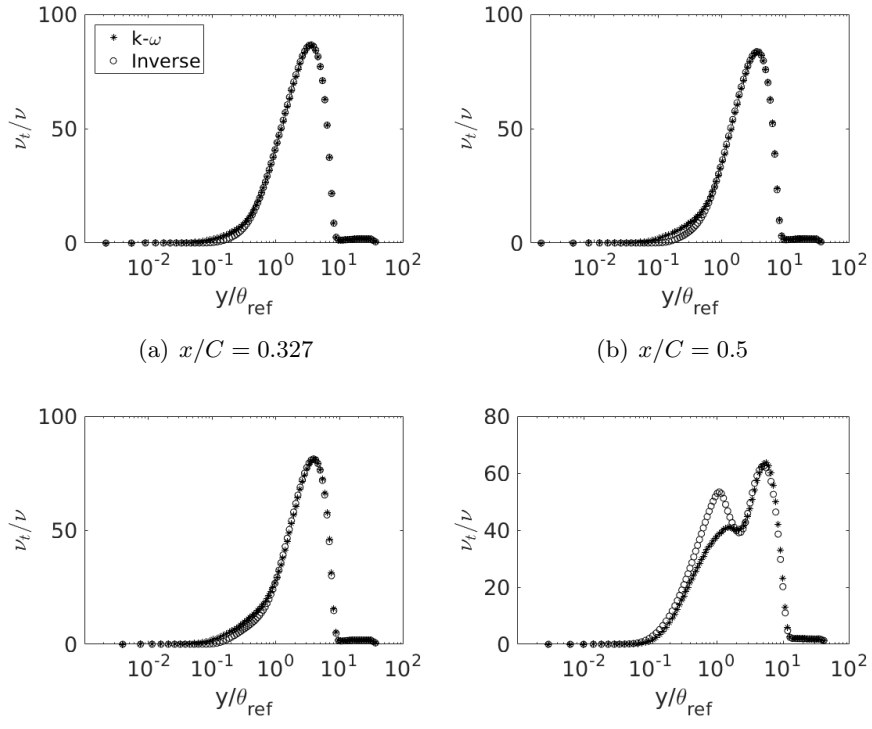
(c) $x/C = 0.655$ (d) $x/C = 1$

Figure 3.31 Comparison of streamwise velocity at various x locations (h20). $*$ base model, $\circ$ optimized, —— LES.



(a) $x/C = 0.327$ (b) $x/C = 0.5$ (c) $x/C = 1$

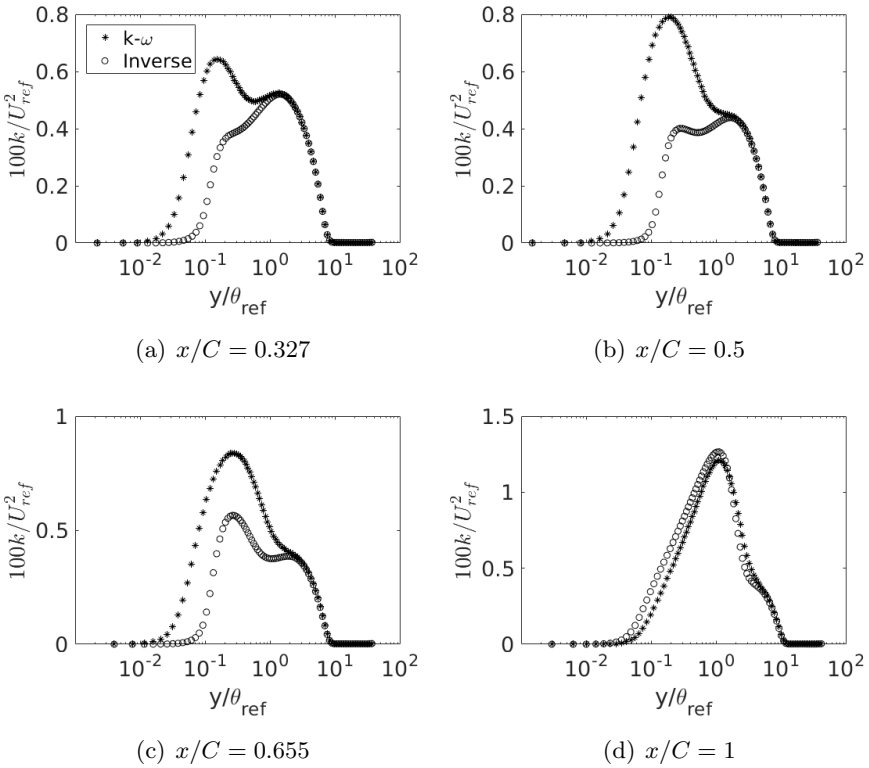
Figure 3.32 correction profiles (β) at various x locations (h20).



(a) $x/C = 0.327$ (b) $x/C = 0.5$

(c) $x/C = 0.655$ (d) $x/C = 1$

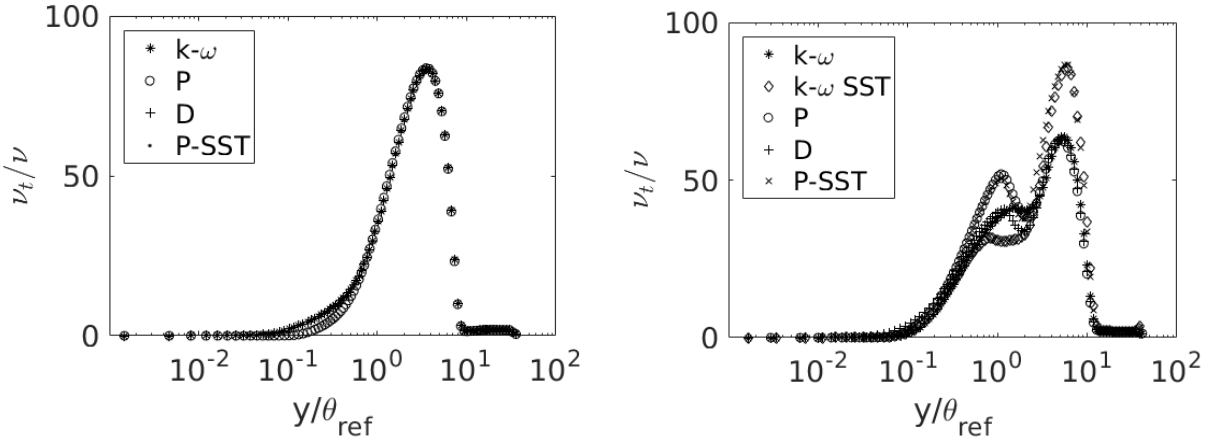
Figure 3.33 Comparison of ν_t at various x locations (h20). $*$ base model, $\circ$ optimized.



(a) $x/C = 0.327$ (b) $x/C = 0.5$

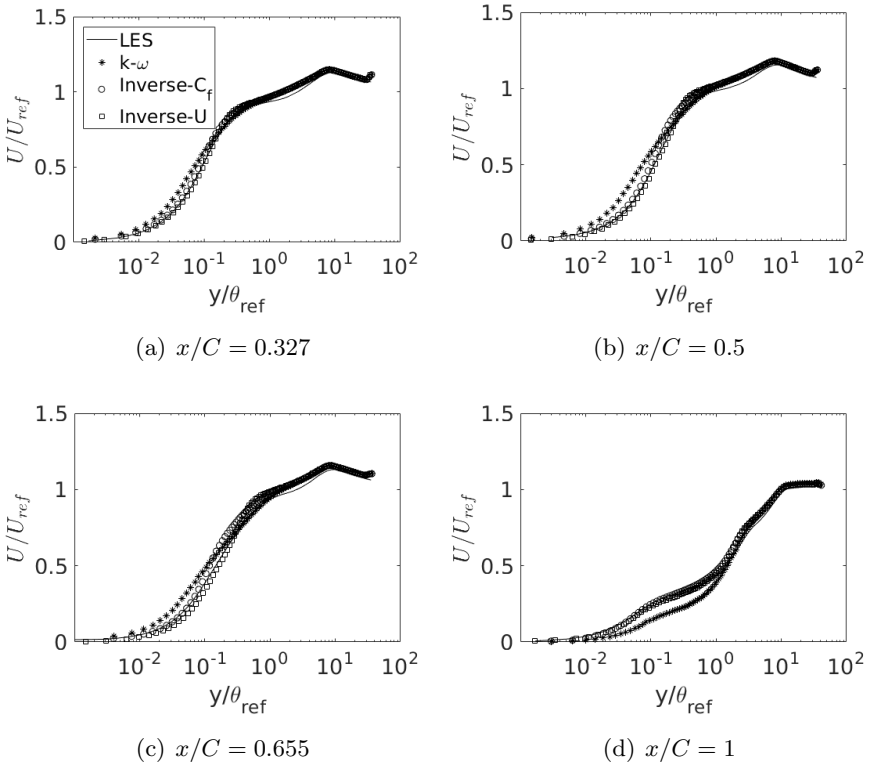
(c) $x/C = 0.655$ (d) $x/C = 1$

Figure 3.34 Comparison of k at various x locations (h20).



(a) $x/C = 0.5$ (b) $x/C = 1$

Figure 3.35 Comparison of inverse solutions with different base models and with the correction factor multiplied to the destruction of ω team instead of production. (h20). $*$: base k-ω model, $\diamond$: base k-ω SST, $\circ$: β multiplied to production of ω of base k-ω model, $+$: β multiplied to destruction of ω of base k-ω model, $\times$: β multiplied to production of ω of base k-ω SST model.



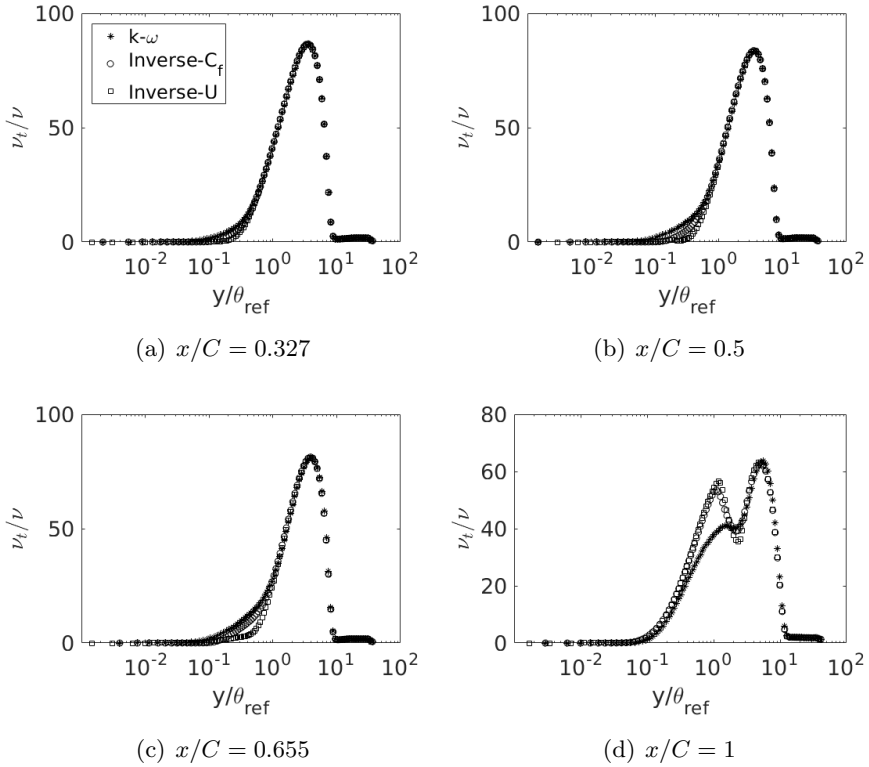
(a) $x/C = 0.327$ (b) $x/C = 0.5$

(c) $x/C = 0.655$ (d) $x/C = 1$

Figure 3.36 Comparison of streamwise velocity at various x locations for inverse solutions with different cost functions (h20).

We regard the resulting eddy viscosity to be extracted data. Fig. 3.33 shows the initial eddy viscosity, corresponding to $\beta = 1$, and the optimized eddy viscosity. It can be seen that most of the difference between the two is near the wall, below $y/\theta_{ref} = 1$ ($y/\theta_{ref} = 1$ corresponds to a $y_+ = 100$). For the favorable pressure gradient side, the baseline model over predicts the eddy viscosity. This trend continues for some distance even after the pressure gradient reverses its sign, suggesting a lag effect. At the end of the bump the baseline model under predicts the eddy viscosity near the wall. One can attribute failures of the baseline $k - \omega$ model largely to erroneous eddy viscosity quite near the wall. Regions where the optimal and baseline viscosity are the same, are insensitive to the optimizer, and these cannot be trusted as extracted data.

The variable k shows a trend similar to the eddy viscosity since they are directly proportional (fig. 3.34). In the FPG section of the bump, k is over predicted by the baseline RANS solution.



(a) $x/C = 0.327$ (b) $x/C = 0.5$

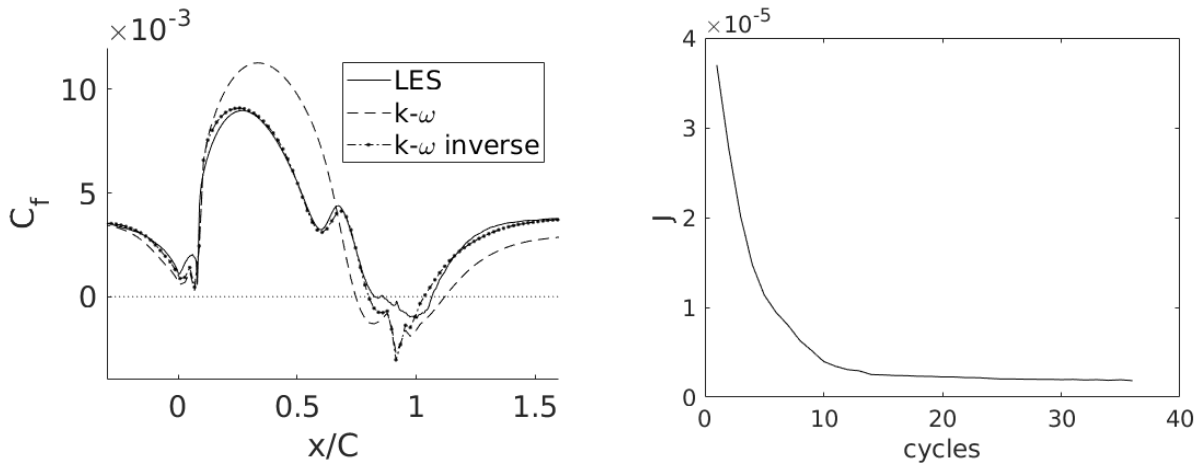
(c) $x/C = 0.655$ (d) $x/C = 1$

Figure 3.37 Comparison of ν_t at various x locations for inverse solutions with different cost functions (h20).

The baseline also predicts a spurious peak near the wall, which does not exist in the optimum solution. These k are quite different from the turbulent kinetic energy computed from the LES solution.

The sensitivity of the inverse solution to the position of the correction multiplier was checked by moving β to the destruction term of the ω equation. Fig. 3.35 shows that the inverse solutions for the eddy viscosity coincide at the top of the bump but differ at the end. This suggests that the optimum is not unique, but nevertheless much different from the RANS solution. Therefore, it was seen that different profiles of eddy viscosity can be considered "correct" and be used to get the actual (LES) velocity field.

The effect of cost function on the inverse solution was examined by using the second cost function as in Eq. 4.3. Fig. 3.36 shows the velocity profile for both the cost functions. A small

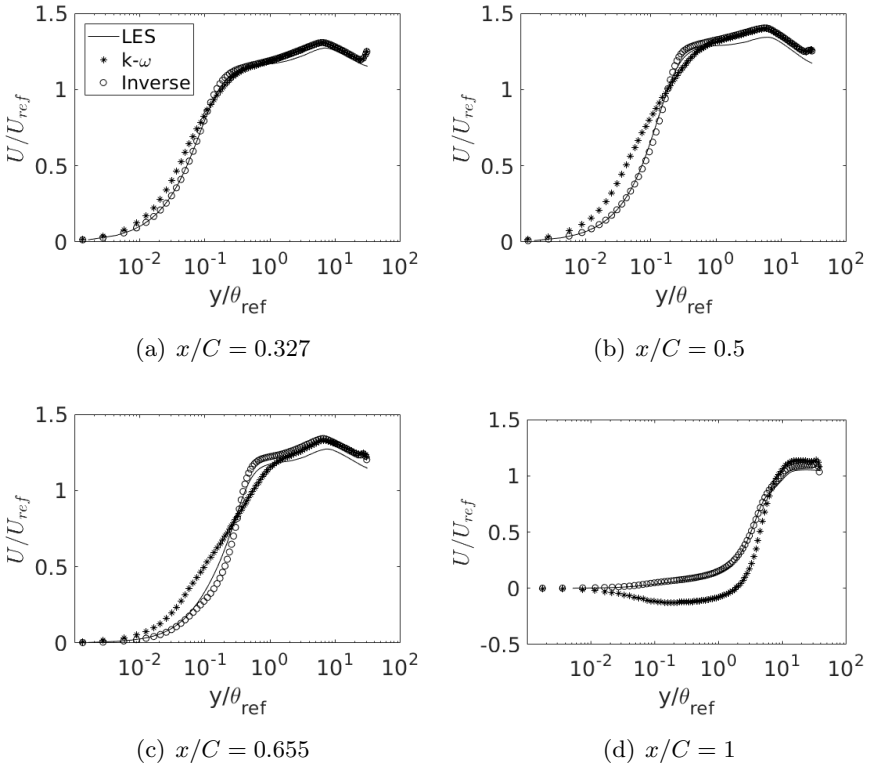


(a) C_f (b) Objective Function

Figure 3.38 C_f of inverse solution compared to LES and RANS solutions and the objective function for the optimization (h38).

difference is visible only at $x/C = 0.655$. The eddy viscosity profiles show some differences (fig. 3.37) suggesting a non-unique optimum eddy viscosity as observed before. Near the end of the bump they become identical. The small values of U near the wall have little effect on the cost-function; that is why inverse wall distance weighting was used in Eq. 4.3. However, the C_f cost-function seems preferable. The C_f cost function is less expensive computationally, since it only requires the calculation of objective function only at the wall, whereas the velocity cost function is calculated at all grid points. However this is not a big difference.

The sensitivity to the base model was tested using the $k-\omega$ SST model, with the corrector term multiplying the production of ω. SST is included in Fig. 3.35. The $k-\omega$ SST model is essentially the $k-\omega$ formulation near wall, and the optimum eddy viscosity solutions overlap. However, dependence on the baseline model is seen away from the wall ($y_+ > 100$), as the optimization is not sensitive to these these locations, so the baseline predictions are not altered. Hence, further away from the wall, the values of ν_t remain the default i.e. same as the base model. Based on this and non-uniqueness it can be inferred that 1) the "correct" velocity profiles can be obtained using multiple ν_t and 2) the ν_t profiles are not very important far from the wall.



(a) $x/C = 0.327$ (b) $x/C = 0.5$

(c) $x/C = 0.655$ (d) $x/C = 1$

Figure 3.39 Comparison of streamwise velocity at various x locations (h38).

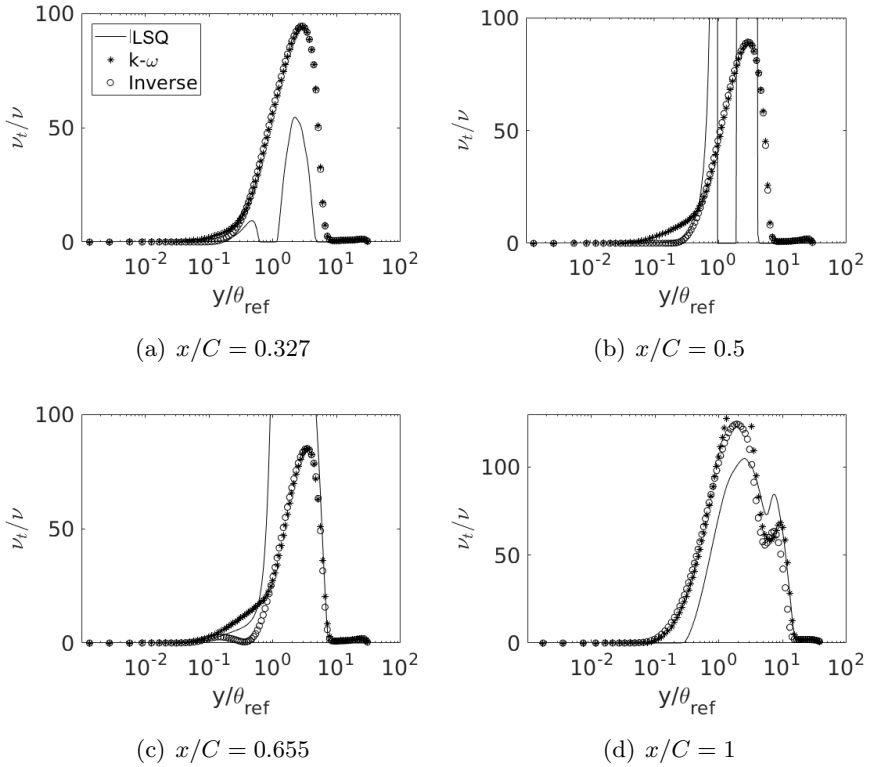
A similar optimization procedure was carried out for h38, and is shown in fig. 3.38. The inverse solution matched almost exactly with the LES data, except near the end of the bump. At this point the gradients become exceedingly small and the optimization does not meet the data as even a large step size in the gradient descent does not change the correction factor appreciably. Elsewhere, it is quite accurate. Fig. 3.39 shows the optimal solution agreeing well with the LES data, again proving that a wall quantity can affect the velocity field further away from the wall.

One might propose a simpler extraction of eddy viscosity. Least squares minimization of the eddy viscosity formula

$$\overline{u_i u}_j = -2\nu_t S_{ij} + \frac{3}{2} k \delta_{ij}$$

with the computed Reynolds stress tensor gives

$$\nu_t = -\frac{\overline{u_i u}_j S_{ij}}{2 S_{ij} S_{ij}} \tag{3.8}$$



(a) $x/C = 0.327$ (b) $x/C = 0.5$

(c) $x/C = 0.655$ (d) $x/C = 1$

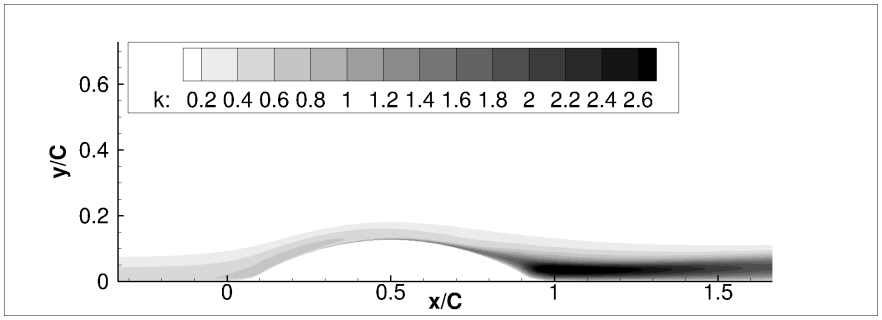
Figure 3.40 Comparison of ν_t at various x locations (h38). $*$ base model, $\circ$ optimized, —— least squares.

Note, however, that there is no connection between least squares and the operational use of eddy viscosity. Fig. 3.40 shows that the eddy viscosity from least squares agrees with the inverse solution near the wall at some locations. However, in general it is not reasonable.

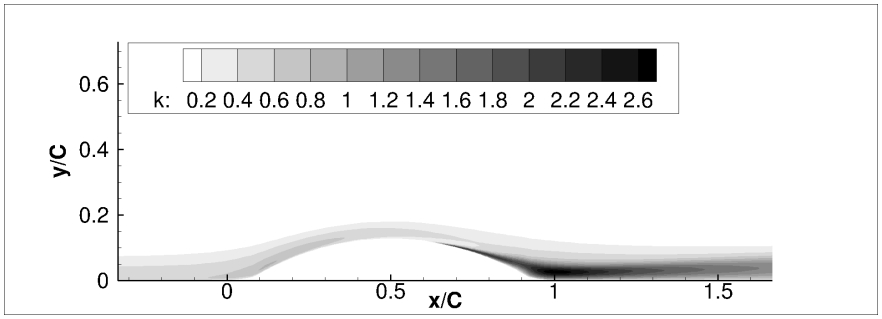
Fig. 3.41 shows the comparison between the k for RANS and inverse solutions: note that k does not represent TKE. It can be seen that after optimization the band of high k extends closer to the wall, similar to the TKE in the LES.

3.5 Conclusion

The present is an empirical study. LES was used to study a set of five bumps, with increasing crest heights. The influence of pressure gradients on the turbulence is brought out through comparisons within this family. The base case, with the lowest bump height, had been studied



(a) RANS



(b) Inverse

Figure 3.41 Comparison of k for RANS and inverse solutions; terms multiplied by $10^2/U_{ref}^2$.

experimentally by Webster et al. (1996) and the LES were validated with these data. Conversely, the LES fills in omissions from laboratory data, providing a more complete picture of the flow field for the reference bump.

The lowest bump had no separation, while, the highest bump produces a marginal separation. The grid requirements are the most severe for the highest bump. To ensure solution accuracy, a convergence study was carried out for that case. For LES, grid refinement studies provide a measure of 'experimental uncertainty'. The grid that was used in the present simulations provides a benchmark degree of accuracy. A database for the bump series is available at Rumsey (2018).

The Reynolds stress data show some intriguing behaviours. All components of the Reynolds stress developed second peaks over the bump indicating the presence of internal layers. A detached layer of high TKE developed in the wake. This is distinct from high TKE in fully separated flow. Even in the marginally separated cases, the band of high TKE leaves the surface well before

separation. The band is due to the mean shear profile created by adverse pressure gradient. Budgets showed that the high TKE near the wall was due a combination of production and turbulent transport.

A plateau in the skin friction develops into a local maximum, as the bump height increases. In combination with the band of high TKE, this is seen to be a consequence of the reaction of the Reynolds stresses to the flow field, and not a direct effect of the pressure distribution.

These, and other effects of strong disequilibrium, such as the absence of a log-layer, prevent data correlation. Hence, the parametric series provides a systematic view of the development of disequilibrium.

Although the focus is on adverse pressure gradient, there is flow acceleration over the front slope of the bump. In that region, turbulent intensity is initially reduced near the wall — increasingly so as the bump height increases. However, the flow remains fully turbulent. Farther from the wall, intensity increases with bump height. There has been some question over whether the flow relaminarizes: it does not, even for the highest bump.

The database was originally created for assessment and development of Reynolds averaged models. It was shown that RANS with $k - \omega$ (or any 2 equation closure model) fails to correctly simulate flow over these bumps. In order to explore the origin of this failure, turbulence variables were extracted by an optimization procedure. The extracted data include the eddy viscosity, ν_t. It showed that current turbulence models fail because their predicted eddy viscosity is erroneous close to the wall. Further away from the wall, the eddy viscosity profiles did not have an appreciable effect on the flow solution.

CHAPTER 4. ZONAL EDDY VISCOSITY MODEL USING MACHINE LEARNING

Adapted from Matai R, Durbin PA (2018) Zonal Eddy Viscosity Models based on Machine Learning. Flow turbulence and combustion Submitted

Abstract

A zonal $k-\omega$ model is constructed, with the zones created by training a decision tree algorithm. The training data are optimized, model coefficient fields. Coefficient data are binned, with each bin assigned a particular coefficient value. The zones are parameterized by training the machine learning model with a local feature set. The features are coordinate invariant flow parameters. It is shown that this model gives superior performance, compared to the base model, in the incompressible adverse pressure gradient (APG) flow test cases. The correction produced by the machine learning algorithm is self-consistent; i.e. once the solution converges, the zones remain fixed.

4.1 Introduction

It has, at times, been suggested that the coefficients of a turbulence closure model might take on different values in pre-set zones. For instance, different coefficients might be provided in free and in wall-bounded shear flows. The idea did not take hold. Recent exploration of the application of machine learning methods to Reynolds averaged closure (RANS) modeling, suggests a new perspective on zonal modeling: decision trees might be trained to identify zones, and assign different model constants to them. We initiate that idea, herein, in the simplest form of training on a single model constant.

The model we have selected is the $k-\omega$ model of Wilcox (2008). This version of $k-\omega$ was selected as the base model because it does not require wall functions, and is not sensitive to free stream conditions. The constant $C_{\omega 1}$ will be given zonal values. To that end, data are needed.

To obtain those data, we adopt the field inversion and machine learning methodology (FIML) of Parish and Duraisamy (2016b). In this approach, the model constant is written as a variable coefficient, $\beta(x)\, C_{\omega 1}$, and β is found by solving an optimization problem. This produces fields of $\beta(x)$. The flow solution provides parameters, like $\nu_T(x)/\nu$, which are supplied, along with β, to the machine learner. The training algorithm develops a functional dependence of β on the parameters. This learned coefficient is inserted into the predictive model.

Singh et al. (2017) applied adaptive boosting (Adaboost), to *correlate* the present coefficient data. Although Adaboost was able to correlate our bump data, it was necessary to use the bump length as a parameter, and the model was not general. Here, we treat the learner as a *classifier*. While that leads to model coefficients that are discontinuous between zones, the gradations are sufficiently small, and the transport equations are sufficiently diffusive, that their solution is smooth.

Combinations of optimization and machine learning have been pursued by other researchers. In an attempt to reduce the modeling error of the $k-\epsilon$ model, Yarlanki et al. (2012) trained a neural network to predict error and used it as an objective function for an optimization algorithm to optimize the model constants. Machine learning (ML), also, has been applied to turbulence modeling by Duraisamy et al. (2015), and to classifying regions of uncertainty by Ling and Templeton (2015). In a conceptual study, Tracey et al. (2015) recreated the source term in the S-A model using a neural network. In that case, the data were created by the S-A model. So, the conceptual result was that a neural network could replace the source term with a learned correlation.

Some of these papers showed that the feature set (non-dimensional parameters) selection is very important. In most cases, the features were not coordinate invariant, and, sometimes, dimensions of the particular geometry were used. That was done because, otherwise, it was not possible to train the learner (correlate the data). Here we select only invariant features, and find that training can succeed.

Ling et al. (2016a) and Ling et al. (2016b) explored incorporating invariance into neural nets and random forests, to predict anisotropy of the Reynolds stress tensor. They tried both an invariant basis representation, and the raw Reynolds stress tensor components. Galilean invariance was embedded into the model by incorporating a multiplicative layer at the end of a deep neural network. This approach showed better *a priori* fit to Reynolds stresses than standard linear, or quadratic eddy viscosity models. It was concluded that an invariant basis yielded better performance and used less computational power to train.

Wu et al. (2017) evaluated two metrics to quantify the similarity between test and training flows. If the flows were very different, the ML algorithms would not perform well. This gives the modeler an *a priori* estimate whether to use a machine learned model. These authors used random forests to correlate Reynolds stress discrepancies. Wall distance based Reynolds number and turbulence intensity were needed to correlate the data; the latter is not Galilean invariant. Wang et al. (2017) reported that when the training flows and prediction flows have different geometries the improvement was marginal.

Weatheritt and Sandberg (2016) presented a way to use Gene Expression Programming to relate the Reynolds anisotropy stresses to an algebric expression based on scalars and tensors. This approach leads to a mathematical expression for the Reynolds anisotropy. Weatheritt and Sandberg (2017) used this method to improve model predictions in an asymmertic diffuser.

It should be recognized that machine learning methods provide a new set of data analysis tools; tools correlate or classify data, and to rank the importance of invariant parameters. The value of those contributions should be recognized, in addition to judging the prospect for machine learning to improve predictive models.

4.2 Methodology

The FIML methodology (Parish and Duraisamy, 2016b) starts by introducing a correction field, $\beta(x)$, into the turbulence model. In the native model $\beta = 1$. The spatial distribution of β is obtained by minimizing a cost function. The cost function introduces data from lab experiment

or high-fidelity computer simulation; the minimization extracts data on β, and on variables of the model. We will use the $k-\omega$ model (Wilcox, 2008)

$$\frac{D\omega}{Dt} = \beta(\boldsymbol{x})\, c_{\omega 1}|S|^2 - c_{\omega 2}\omega^2 + \nabla\cdot(\nu + \sigma_\omega k/\omega)\nabla\omega + (\sigma_d/\omega)\nabla k\cdot\nabla\omega \tag{4.1}$$

While the Wilcox (2008) model has some advantages, FIML can just as easily be applied to $k-\omega$ SST, Wilcox (1988) and many other models. Notice that the correction field variable, β, multiplies the production term of the ω equation. Matai and Durbin (2018) show the method is equally effective if β multiplies the destruction term.

For future reference, we note that

$$\sigma_d = 0 \text{ if } \quad \nabla k\cdot\nabla\omega \le 0, \text{ and } \sigma_d = 1/8 \text{ otherwise} \tag{4.2}$$

The cost function, J, will be the skin friction coefficient, with data provided from the LES of Matai and Durbin (2018). Thus, $\beta(x,y)$ minimizes

$$J(\beta) = \sum_i (C_f^{RANS}(\beta) - C_f^{LES})_i^2 \Delta\ell_i \tag{4.3}$$

where the summation is over the surface discretization. (Matai and Durbin (2018) show very similar results when the cost function uses $U(x,y)$; using C_f gave a slightly better fit.) A gradient descent algorithm was used to minimize J. The gradient $(\partial J/\partial\beta)$ was found by solving an adjoint equation: the method and its application to the present LES data is described in Singh et al. (2017).

The correction field can be regarded as data, that were extracted from a particular experiment. To produce a predictive model, $\beta(x,y)$ must be correlated with an invariant set of parameters (a.k.a., features: see table 4.5). A machine learning (ML) model establishes a functional relation between field variables and the correction factor. Call the feature set $\{\eta_i\}$: the ML model is $\beta(\eta)$. In the present paper, the ML model is (a bag of) decision trees. The trees partition the flow domain into zones, in each of which a particular value of β is used.

To that end, the correction field data were divided into bins. These bins served as the classes for the decision tree classifier. The value and size of bins was decided after experimenting with different values and observing how the training cases were affected. The sensitivity to bin size and

Table 4.1 Bins for classification of zones

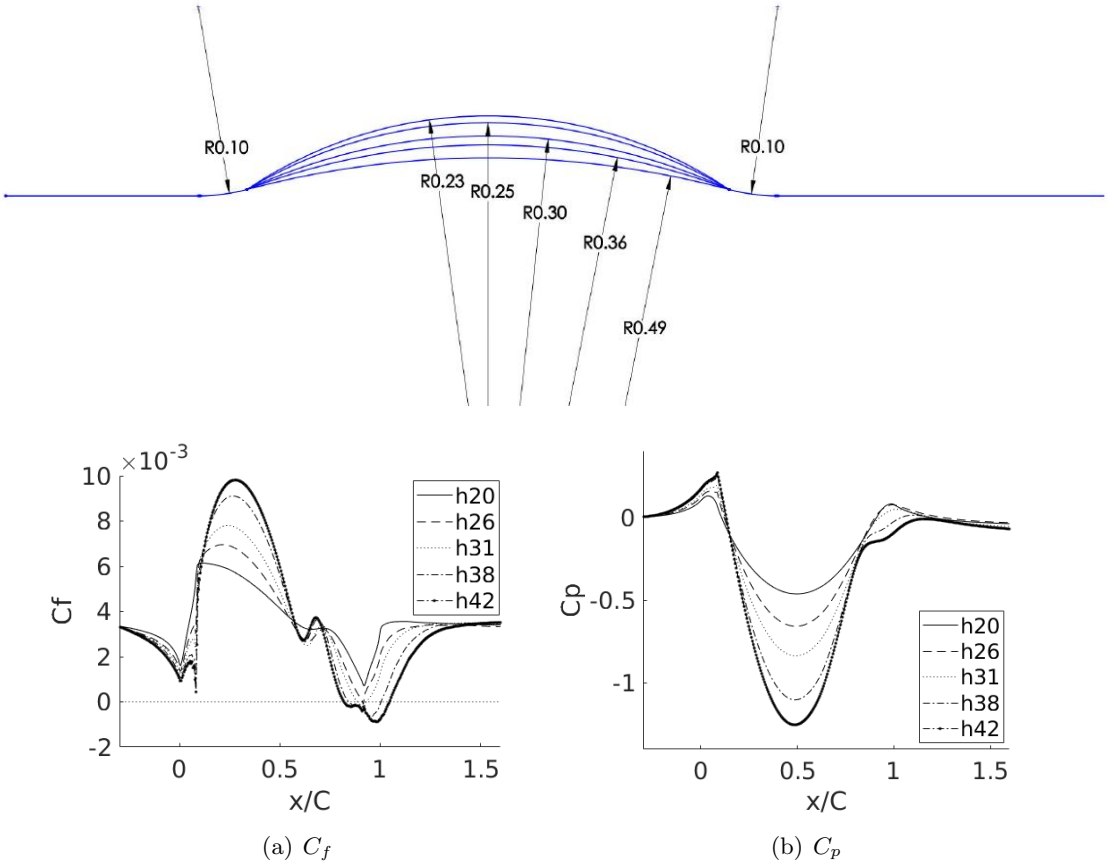
Bin Number	*Bin*	β	*Population of Samples in bin*
1	$0.2 >$	0.1	284
2	$0.2 \leq < 0.4$	0.2	327
3	$0.4 \leq < 0.6$	0.4	609
4	$0.6 \leq < 0.95$	0.6	3215
5	$0.95 \leq < 0.97$	0.8	696
6	$0.97 \leq < 1.275$	1	2160
7	$1.275 \leq < 2$	1.75	772
8	$2 \leq < 2.5$	2.25	316
9	$2.5 \leq < 3$	2.75	219
10	$3 \leq$	3.25	256

class values is discussed in section 4.2.1. The bins for the final model are listed in table 4.1. The bag of trees classifier maps the feature set into these bins. Several other methods such as pure decision trees, support vector machines, and adaBoostM1 were considered, but the bag of trees yielded the best results.

Another pre-processing step must be mentioned. Over most of the domain $\beta = 1$. Non-unity corrections were concentrated near the wall. Therefore, before training the classifier, data points with $0.98 < \beta < 1.05$ were removed. This pre-processing was required to prevent the classifier from being overwhelmed by irrelevant data. As the optimizer made corrections primarily near the wall, training data were restricted to that region. The restriction was achieved by using the definition of σ_d in equation 4.2 to identify the region of interest. The closure model uses σ_d to transition between $k-\omega$ behavior near the wall, and $k-\varepsilon$ behavior in the outer part of a boundary layer. So, in the $k-\omega$ region, the optimized β were used; elsewhere $\beta = 1$.

All the RANS calculations were carried out in OpenFOAM (Weller et al., 1998) and the machine learning algorithm was trained in Matlab. The $k-\omega$-ML model will be referred to as $k-\omega$-*mod*. The modified model calculated the value of the feature set, then queried the trained Matlab ML algorithm, to obtain a correction factor at every grid point. This process was iterated until

convergence; i.e. changes in both the correction factor, and the full RANS solution, were below a tolerance. The hyper-parameters of the bag of trees are those in table 4.2: it contains 40 trees.

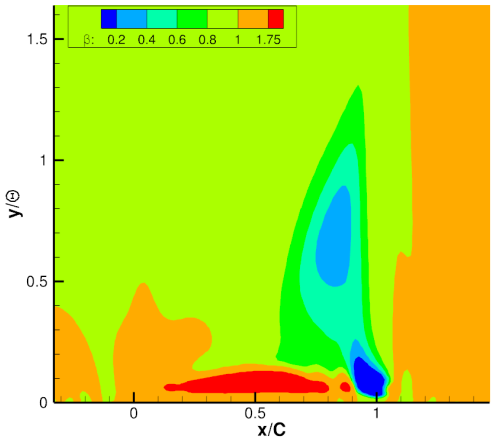


(a) C_f (b) C_p

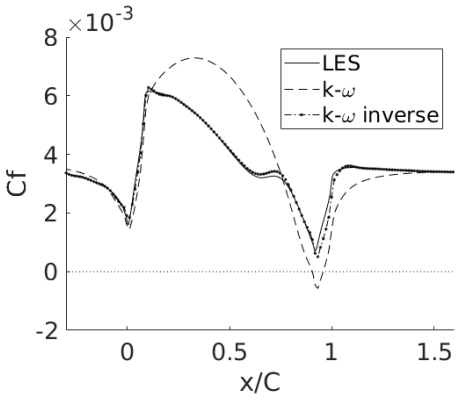
Figure 4.1 Comparison of skin friction coefficient and pressure for different bump crest heights

The training and test data come from LES of a set of bumps of increasing height, as in figure 4.1 (Matai and Durbin, 2018). The bumps are labelled by their height in mm; the bump length is $C = 305mm$. A full set of mean velocities and Reynolds stresses is available, but only the C_f, shown in the figure, were used to extract β by optimization.

The model was trained on bumps h20 and h38. Per figure 4.1, there is a short region of adverse pressure gradient (APG) before the start of the bump, followed by favorable pressure gradient (FPG) on the windward side, followed by adverse pressure gradient on the leeward side, and a recovery region over the flat section after the bump, $x/C > 1$. Case h20 provides a flow just



(a) β for h20



(b) C_f comparison for base and optimized (inverse) model

Figure 4.2 Comparison of skin friction coefficient and pressure for different bump crest heights

about to separate, and h38 provides a flow that has a small separation bubble. A fully developed boundary layer, having $Re_\theta = 2{,}500$, was imposed at the inlet, in all cases.

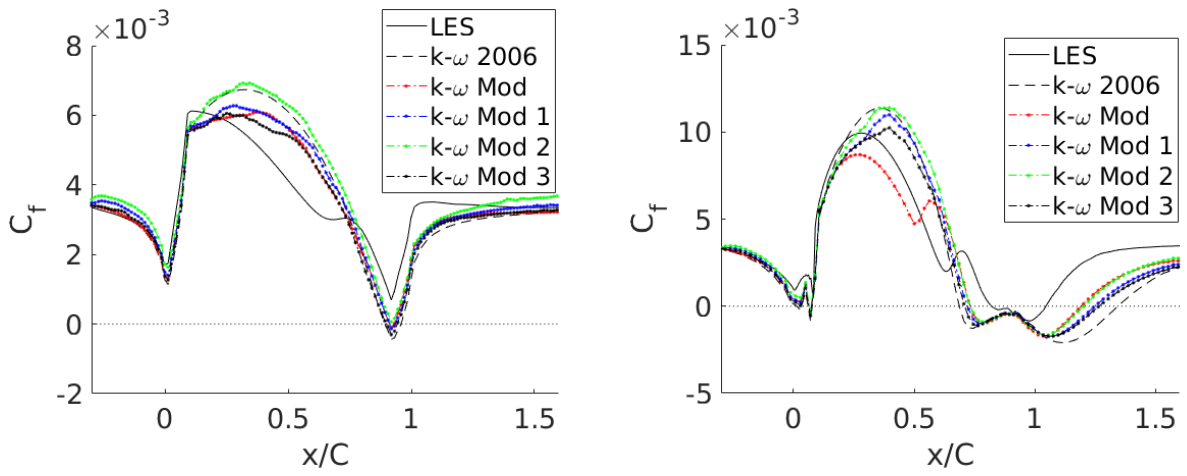
Table 4.2 hyper parameters for bag of trees

number of splits	900
number of trees	40
splitting criteria	Gini's diversity index

4.2.1 Sensitivity Analysis

4.2.1.1 Sensitivity to bin width

In this method, the size of the bins, and the correction values, β_i, assigned to each bin, are tuned. There are no set guidelines for this. Intuitively, the value of the correction value should lie within, or be close to, the bin. Multiple such values were experimented upon. This section explores the effect of the number of bins on one training flow (h20) and one test flow (h42).



(a) $C_f h20$ (b) $C_f h42$

Figure 4.3 Sensitivity of C_f to bin width on training (h20) and test (h42) data.

First, the correction value was distributed among 14 bins with each bin having a width of 0.2, and having a correction value equal to the lower end of the bin, as in table 4.3. This is referred to as model 1. Second, the correction value was distributed among 8 bins with a width of 0.4, again having a correction value equal to the lower end of the bin, as shown in table 4.4. This is referred to as model 2. Another model with bin width of 0.1 was also tested (model 3) and it can be seen a finer bin doesn't necessarily lead to better results.

As seen in fig. 4.3, both models performed better than the base model. They have similar performance in the recovery region, but more bins results in better performance over the bump. This is due to the correction being equal to the lowest value of the bin, and thus biasing the correction towards lower values. In fig. 4.7 correction values are less than unity around the recovery region, and unity or greater, elsewhere over the bump.

In the h20 plot of fig. 4.3, C_f is over predicted in $x < 0$, leading up to the bump. This is a region of mild APG/ZPG and this region should have a correction factor close to unity; but it is actually much lower. From this it can be concluded that the models would over predict skin friction over a ZPG flat plate. (Indeed, this was seen in a ZPG flat plate computation). Similarly, model 2 over predicts C_f in the recovery region, which is mild FPG/ZPG, again indicating that a lower

Table 4.3 Bins for classification of zones for modified model 1

Bin Number	*Bin*	β	*Population of Samples in bin*
1	$0.2 >$	0.1	284
2	$0.2 \leq < 0.4$	0.2	327
3	$0.4 \leq < 0.6$	0.4	609
4	$0.6 \leq < 0.8$	0.6	1250
5	$0.8 \leq < 1.0$	0.8	3316
6	$1 \leq < 1.2$	1.0	1294
7	$1.2 \leq < 1.4$	1.2	449
8	$1.4 \leq < 1.6$	1.4	244
9	$1.6 \leq < 1.8$	1.6	133
10	$1.8 \leq < 2.0$	1.8	157
11	$2.0 \leq < 2.2$	2.0	117
12	$2.2 \leq < 2.4$	2.2	159
13	$2.4 \leq < 2.6$	2.4	70
14	$2.6 \leq$	3.0	445

Table 4.4 Bins for classification of zones for modified model 2

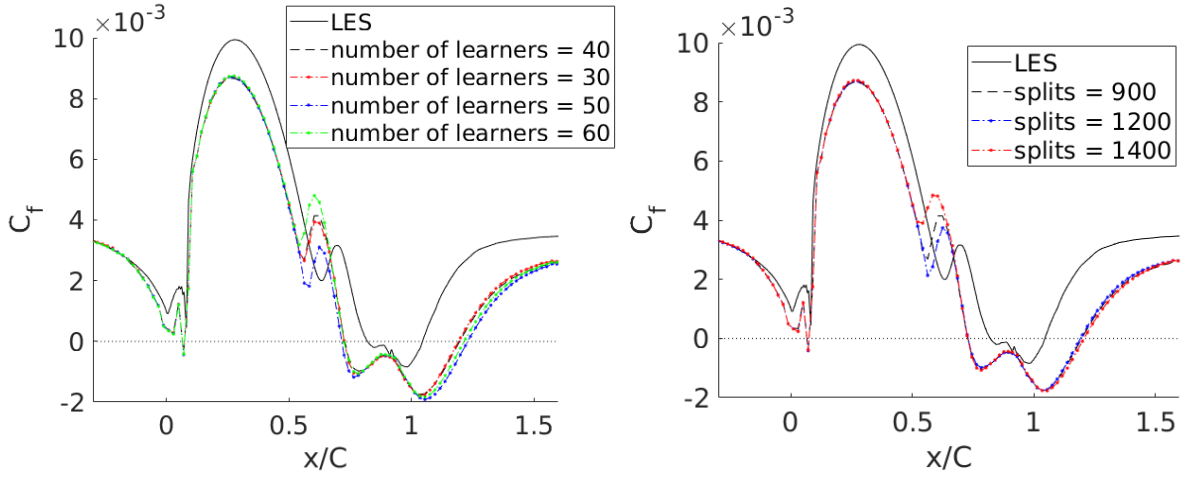
Bin Number	*Bin*	β	*Population of Samples in bin*
1	$0.2 >$	0.1	284
2	$0.2 \leq < 0.6$	0.4	936
3	$0.6 \leq < 1.0$	0.6	4566
4	$1.0 \leq < 1.4$	1.0	1743
5	$1.4 \leq < 1.8$	1.4	377
6	$1.8 \leq < 2.2$	1.8	274
7	$2.2 \leq < 2.6$	2.2	229
8	$2.6 \leq$	3.0	445

than required correction was generated. In this region, the base model does under predict the C_f, therefore a mild, less than unity correction is required; but since model 2 has larger bins, it outputs a lower than required correction. One can also conclude that model 2 would over predict C_f in a channel flow. From the above experiment it can be seen that the bin value and width near $\beta = 1$ is important to correctly predict canonical wall bounded turbulent flows. These experiments, and many similar, led to the bins of the final model (table 4.1). They were designed to have a value, close, or equal, to unity, from 0.95 to 1.275. The rest of the bins were allowed to be larger. This provides some improvement for h20, and more noticeable improvement for h42. The bins may not be optimal but, as will be seen, they work well.

4.2.1.2 Sensitivity to number of learners and splits

Increasing the number of learners can increase the accuracy. Thus a number of cases, ranging from number of learners = 30-60, were tested. It was seen that increasing this number beyond 40 slightly deteriorated the performance in the recovery region as seen in figure 4.4 and thus the number of learners was set to 40. Increasing the maximum number of splits (and therefore the depth) of each tree can also increase the accuracy (but may lead to over fitting). It was seen that increasing this number beyond 900 splits did not yield any benefit in this test case.

Finally to ensure that the model was reproducible, multiple instances of a model with the hyper parameters mentioned in table 4.2 were trained and tested. It was seen the major difference lay in the reproduction of the sudden increase in C_f which is a uncommon feature in most flows. This may be due to the randomness associated with selecting data for the training of each tree (the data for each tree is selected randomly with replacement). Other than that the reattachment length remained constant in all models and the models also gave similar results in the FPG region of the flow. The results are summarized in figure 4.5.



(a) number of learners (b) number of learners

Figure 4.4 Sensitivity of C_f to number of learners and number of splits (h42).

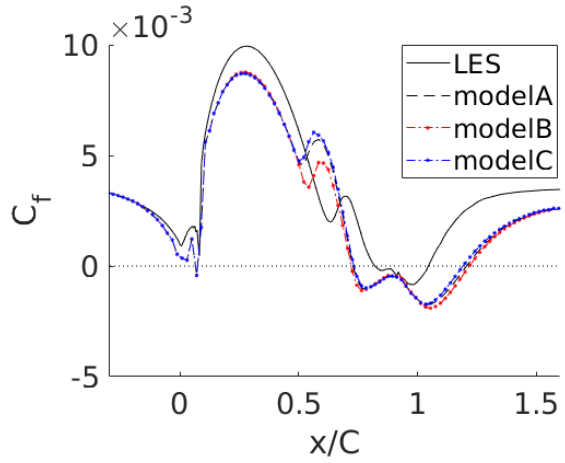


Figure 4.5 Models trained on same data with same hyper parameters (h42).

Table 4.5 Input Feature set. Note that in OpenFOAM incompressible solvers, p/ρ is computed instead of pressure and hence p here has the units $length^2/Time^2$

#	*description*	*feature*	*normalized feature*
1	magnitude of pressure hessian	$\|\frac{\partial^2 p}{\partial x_i \partial x_j}\|$	$\frac{\|\frac{\partial^2 p}{\partial x_i \partial x_j}\|}{\omega^2 + \|\frac{\partial^2 p}{\partial x_i \partial x_j}\|}$
2	pressure velocity gradient	$\partial_i p S_{ij} \partial_j p$	$\frac{\frac{\partial p}{\partial x_i} S_{ij} \frac{\partial p}{\partial x_j}}{k\omega^3 + \frac{\partial p}{\partial x_i} S_{ij} \frac{\partial p}{\partial x_j}}$
3	turbulent viscosity	ν_t	$\frac{\nu_t}{\nu + \nu_t}$
4	magnitude of velocity gradient tensor	$\|S\|$	$\frac{\|S\|}{\frac{\nu\omega^2}{k} + \|S\|}$
5	wall distance	d	$\frac{\sqrt{k}d}{\nu}$
6	Q criterion	$\|R\|^2 - \|S\|^2$	$\frac{\|R\|^2 - \|S\|^2}{\|R\|^2 + \|S\|^2}$
7	time scale ratio	$\|S\|$	$\frac{\|S\|}{\omega + \|S\|}$

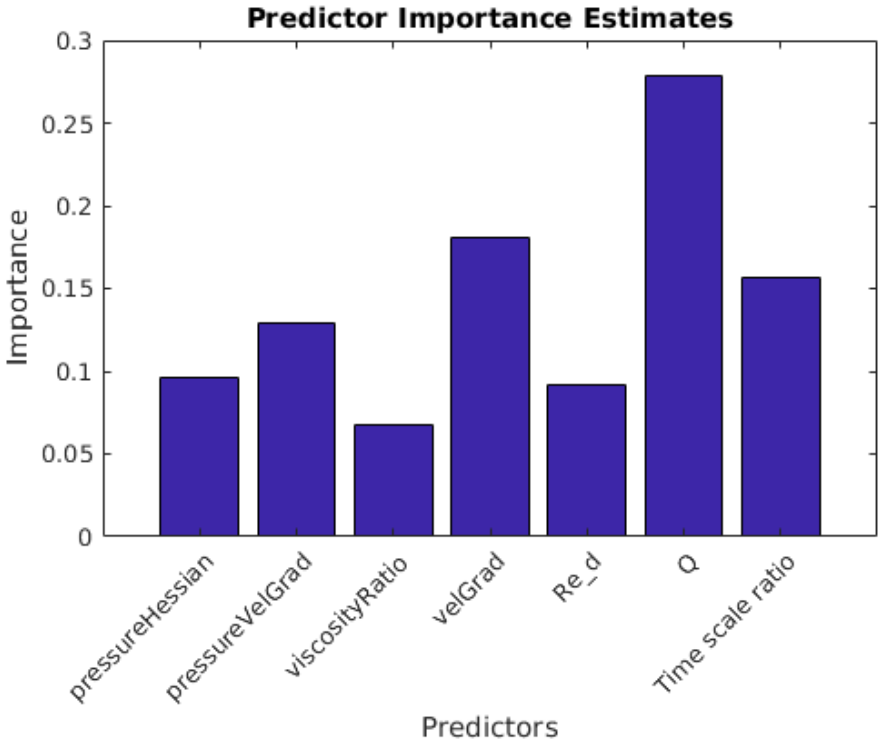


Figure 4.6 Estimated importance of features.

4.2.2 Feature Set Formulation

The features are summarized in table 4.5; they were selected largely by looking at contours of quantities that showed a high visual correlation with the β field. The quantities, ν_t/ν and $\sqrt{k}d/\nu$, that are used in existing turbulence models, also were included. It is inescapable that the selection of features introduces human judgement; indeed, any combination of features could be used, so there is an infinite set from which to choose. Having contours of LES fields is a boon.

Special care was taken to ensure that the ML classifier has access to pressure gradient information, since regions of high APG/FPG result in incorrect predictions by current models. All the features are Galilean invariant, and none of them depend on geometric quantities or wall quantities, further generalizing the model, and making it easier to implement in an unstructured solver. Finally, all the quantities are invariant under rotation. As in Ling and Templeton (2015), aside from $\sqrt{k}d/\nu$, the features were non-dimensionalized in the form $|A|/(|A| + |B|)$ to constrain them to lie in [-1,1].

Fig. 4.6 shows the importance of features, as determined by the (Matlab) decision tree algorithm. Q is the most important feature; $Q > 0$ identifies where rotation is greater than strain. Velocity gradient, pressure-velocity-gradient, and time scale ratio were also important features. It is interesting that the wall distance-based Reynolds number and the turbulent Reynolds number (viscosity ratio) are the least important features. This may be due to training only on filtered points near the wall — as explained earlier.

4.3 Results

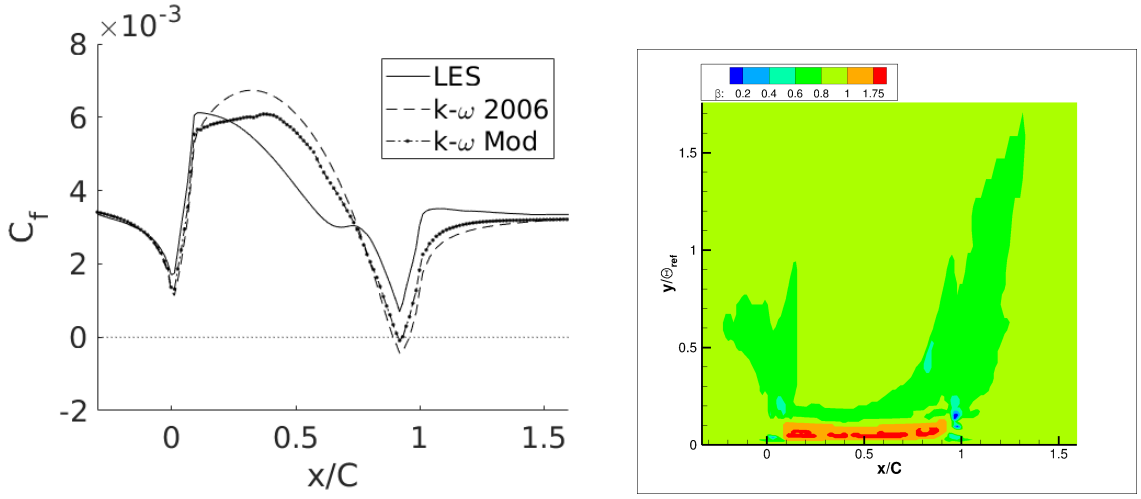
4.3.1 Training Cases

The two cases for training were h20 (with no separation) and h38 (with mild separation) of Fig. 4.1. These two cases were chosen so that the ML algorithm would 'know' the difference between APG flows that separate and those that don't.

Table 4.6 Grid characteristics

flow	Nx	Ny	cells
h20/38/42	89	99	8811
channel	64	64	4096
ZPG BL	69	100	6900
Back Step	112	152/200	15872
Obi Diffuser	362	140	50680
NASA Hump	816	216	176256

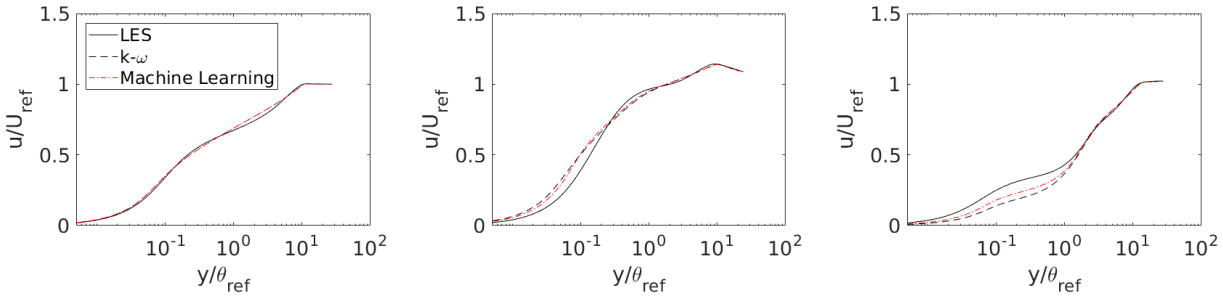
A separate ZPG flat plate computation provided an inlet boundary layer with the required inflow momentum thickness and Reynolds number. For the bump computations, the velocity profile from the ZPG boundary layer was prescribed at the inlet and a zero gradient at the outlet. The first grid point was maintained at $y_+ < 1$ on the bottom wall and the top was assigned a ZPG condition for all the variables. The grid details are given in Table 4.6.



(a) C_f (b) Zones

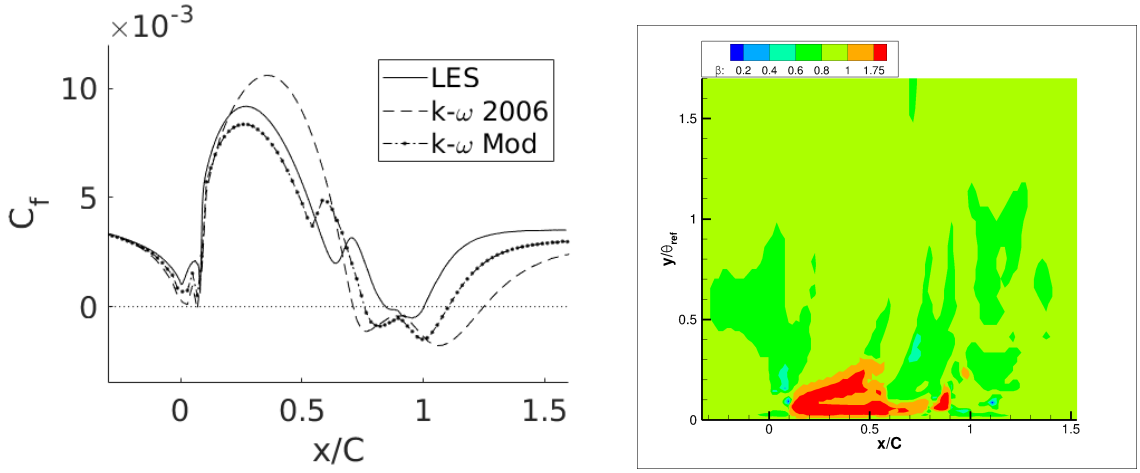
Figure 4.7 Skin Friction Coefficient and zones for h20.

The ML model makes corrections close to the wall, with the production enhancing zones ($\beta > 1$) nearer to the surface than the production damping zones ($\beta < 1$). Also, $\beta < 1$ in the later part of



(a) Inlet (b) $x/C = 0.6$ (c) $x/C = 1$

Figure 4.8 x-velocity at various x locations for h20.



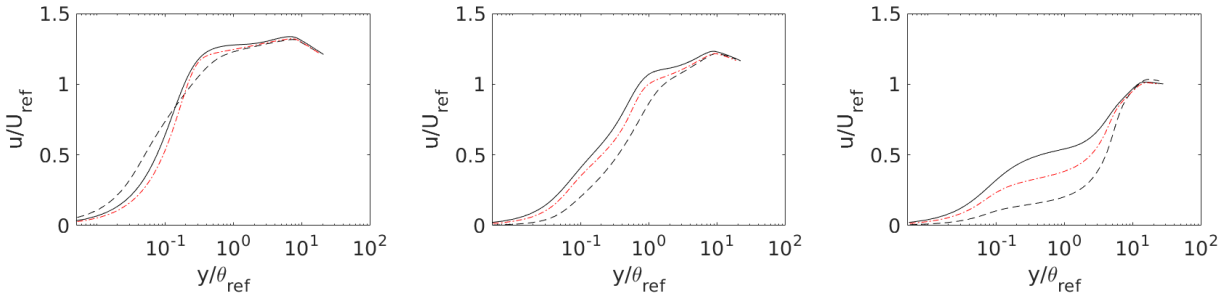
(a) C_f (b) Zones

Figure 4.9 Skin Friction Coefficient and zones for h38.

the APG region, and in the recovery region; while $\beta > 1$ in the FPG and parts of the APG region — as shown in Fig. 4.7.

The base model predicted separation in the h20 case, although there is none (fig. 4.1). It was observed that the $k - \omega$-*mod* reduced the separation, but did not eliminate it completely. The velocity profiles also improved at some of the stations (Fig. 4.8) and so did the recovery of the skin friction coefficient (C_f: Fig. 4.7).

The base model predicted a separation bubble for h38 which was much larger than the LES data. It can be seen that $k-\omega$-*mod* reduced the predicted separation by 33%. The velocity profiles were improved at some of the stations (Fig. 4.10) and so was the recovery of the skin friction



(a) $x/C = 0.5$ (b) $x/C = 0.7$ (c) $x/C = 1.33$

Figure 4.10 x-velocity at various x locations for h38.

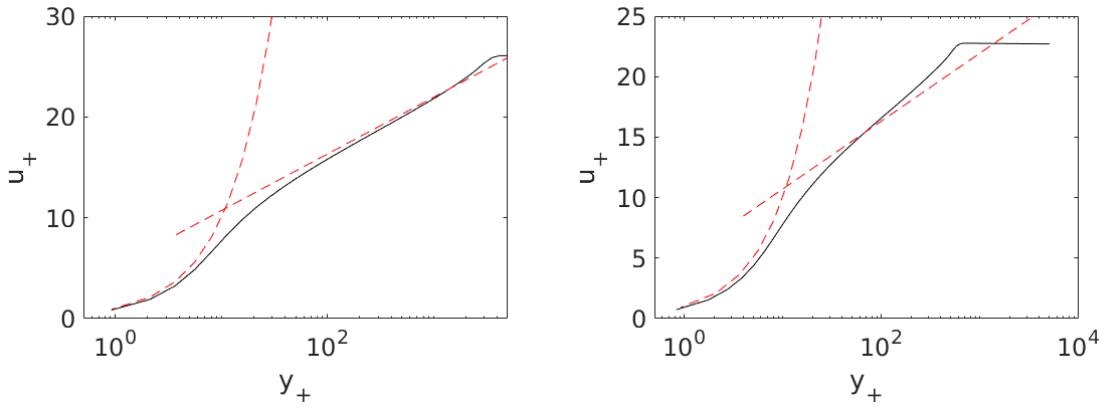
coefficient (C_f: Fig. 4.9). Similarly to h20, $\beta > 1$ in FPG and parts of APG, and $\beta < 1$ in the recovery region and parts of APG. In general, $\beta > 1$ results in a reduction of eddy viscosity (ν_t) which in-turn causes a reduction in wall shear stress (τ_w). Also it was seen that $\beta = 1$ before the bump (mild APG/ZPG) and in the mild FPG after the bump. This was by design, as it ensured that the new model did not alter flat plate solution (ZPG) and channel flow (FPG) solutions, as they are predicted correctly by the base model.

4.3.2 Test Cases

4.3.2.1 Channel Flow and Flat Plate

Flow in a channel and over a flat plate with a zero pressure gradient are standard cases on which a new turbulence model is tested. Since the $k-\omega$ model was used, $y_+ < 1$ was necessary at the walls. Table 4.6 shows the grid characteristics for channel flow and ZPG boundary layer.

Channel flow was computed at $Re_\tau = 5,500$. The velocity profile agrees very well with the log law (Fig. 4.11) and the skin friction is the same as predicted by the base model. Thus the ML model made no significant modification. A similar observation was made for flat plate boundary layer at $Re_\theta = 1,200$, as shown in Fig. 4.11. The log law was reproduced exactly as it was by the base model. This was a satisfactory result since the base model is calibrated to produce the correct log-law for ZPG flat plate and channel flows.



(a) channel flow (b) flat plate

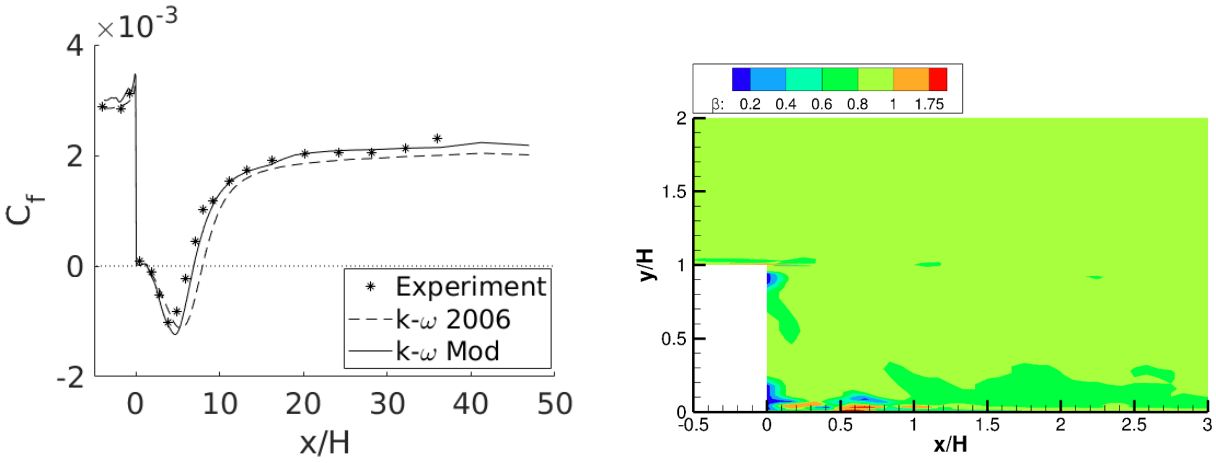
Figure 4.11 velocity profile for channel flow at $Re_\tau = 5,500$ and flat plate at $Re_\theta = 1,200$.

4.3.2.2 Back Step

The backstep (Driver and Seegmiller, 1985) is another standard test case for turbulence models. It is particularly interesting, because it tests the ML model's ability to predict a reasonable β when encountering an abrupt change in flow geometry. This case includes the processes of reattachment and recovery, which are important flow features to capture. The inlet $Re_\theta = 5,000$ is twice the Re_θ of the training cases. The grid details are given in Table 4.6.

An inlet velocity profile was prescribed and the zero gradient BC was used for velocity at the outlet. A separate channel flow computation was used to generate the inlet velocity profile. A reference pressure of zero was specified at the outlet and the inlet was set to ZPG. A $y_+ < 1$ was used near wall.

Although the original model performs reasonably well on this backstep, it over predicts the separation bubble. The ML augmented model reduces this error (Fig. 4.12). The experimental data did not extend close to the wall and the outer flow predicted by both the models was similar, as shown in Fig. 4.13. Fig. 4.12 shows the different zones predicted by the bag of trees. The non-unity zones are concentrated near wall, where most of the correction is needed.



(a) C_f (b) Zones

Figure 4.12 Back Step C_f and zones.

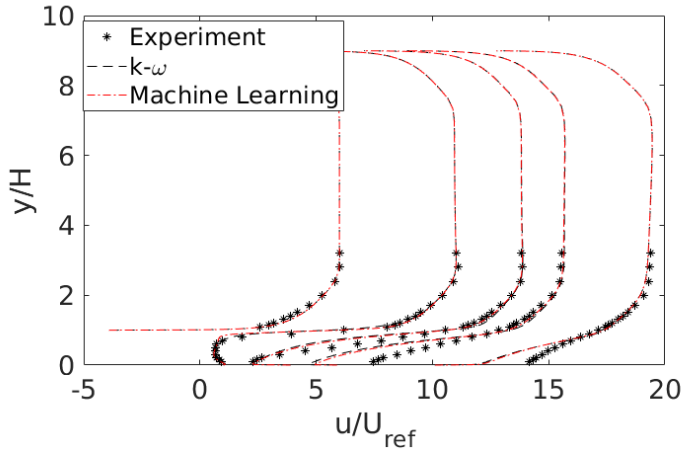
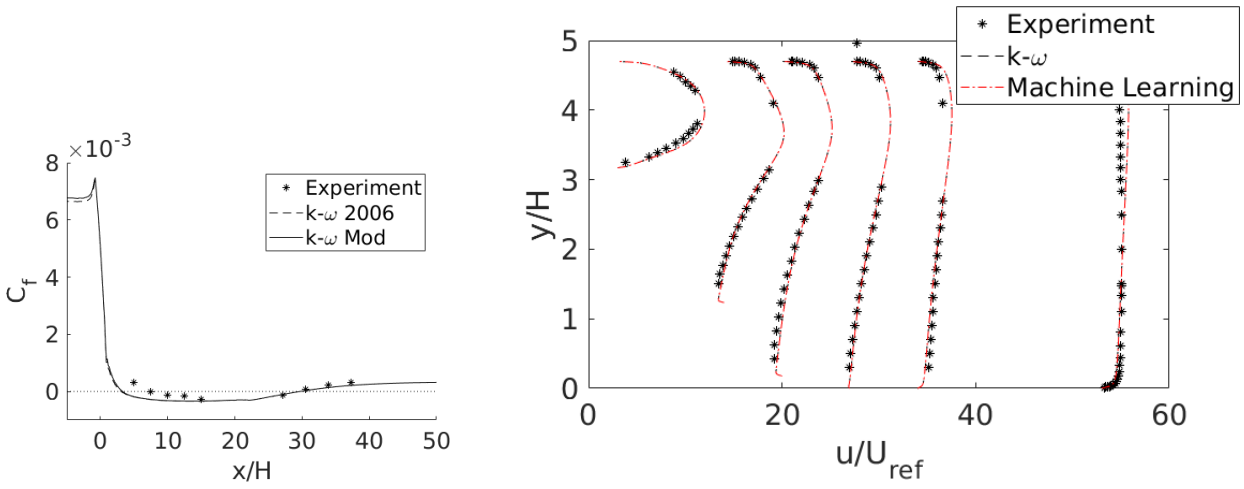


Figure 4.13 Back Step experimental data compared with base model and ML model..



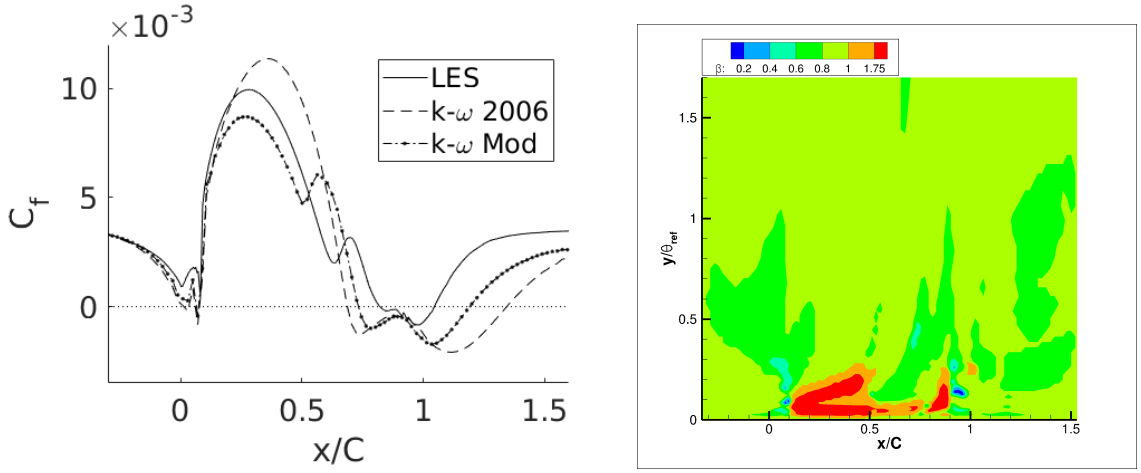
(a) C_f (b) Velocity

Figure 4.14 Obi diffuser experimental data compared with base model and ML model.

4.3.2.3 Obi Diffuser

The Obi diffuser (Obi et al., 1993) is a 2D diffuser, with a relatively small separation bubble. It provides an example of a flow with multiple changes in the sign of the pressure gradient, that is different from the bump series. The flow enters the diffuser from a fully-developed channel (FPG) and this turns to a APG in the diffuser section. The inlet Reynolds number based on channel center line velocity and channel height was 20,000. The fine grid, which had $y_+ \sim 1$, was used with the inlet length reduced to ensure that the velocity at the inlet station were as required. A fully developed velocity profile was specified at the inlet and a zero gradient BC was used at the outlet for the velocity BCs. A reference pressure of zero was specified at the outlet.

The original model does very well on this geometry and the ML model did not make significant changes, as shown in Fig. 4.14. The C_f and velocity profiles are practically identical. As in the channel flow and ZPG boundary layer computations, this showed that the ML algorithm would not make significant changes in cases where the original model was already preforming well even in APG regions.



(a) C_f (b) Zones

Figure 4.15 h42 skin friction and zones.

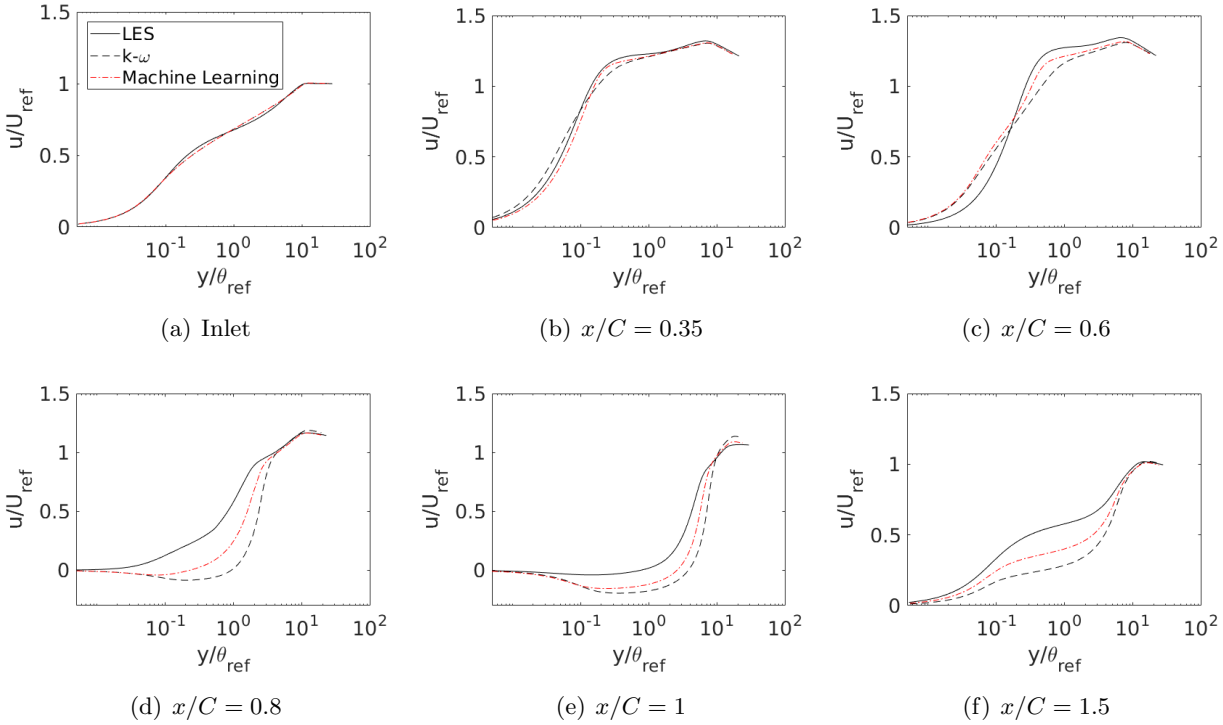
4.3.2.4 h42

Bump h42 is the highest bump in the series (Fig. 4.1); it has the same inlet conditions as h20 and h38. This bump has a slightly larger separation bubble than h38. Its size is over-predicted by the base model. The $k-\omega$ *mod* reduced the bubble and, thus, improved the model predictions. Similar to h38 and h20, the FPG and parts of APG had enhanced production of ω and recovery and parts of APG had suppressed production.

Figure4.16 shows that the ML augmented model made some major changes to the velocity profiles in the APG and the recovery region and the resulting velocity profiles were much closer to the LES profiles. The location of the omega increasing ($\beta > 1$) and decreasing ($\beta < 1$) zones remained similar to the other bumps but the increased in size.

4.3.2.5 NASA Hump

The NASA hump is a 2D hump in a channel, as described in Naughton et al. (2006). This case is similar to the series of bumps, fig. 4.1, in the sense that it also has a APG separation; but it is different, as it has a concave APG compared to the abrupt APG (backstep) and convex APG (bump



(a) Inlet (b) $x/C = 0.35$ (c) $x/C = 0.6$

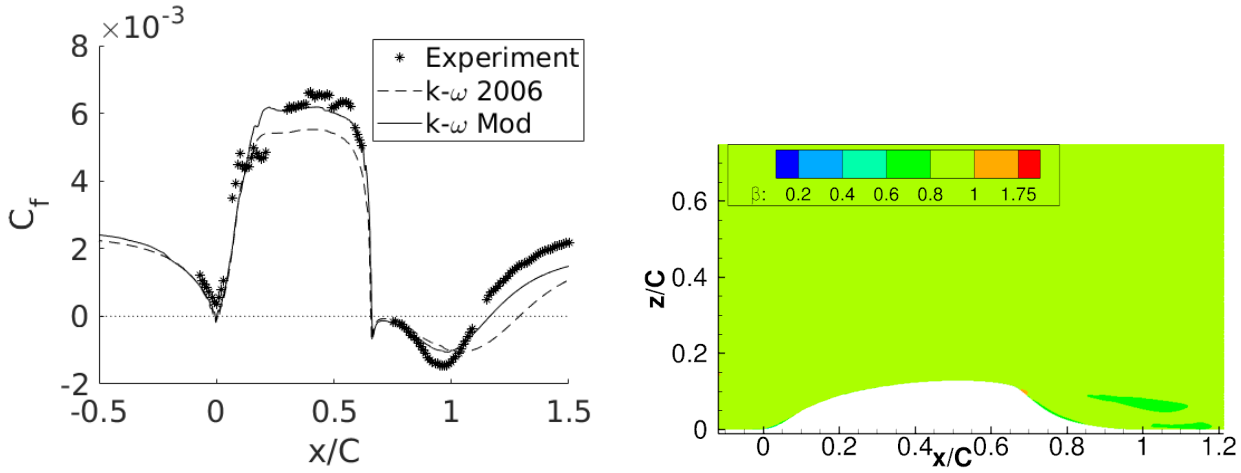
(d) $x/C = 0.8$ (e) $x/C = 1$ (f) $x/C = 1.5$

Figure 4.16 x-velocity at various x locations for h42.

series). The original case also had a plenum at the beginning of APG but in this computation the plenum was not present.

The second finest grid described in Greenblatt (2018) was used. The long inlet section allowed for a fully developed boundary, that matched the experimental thickness at $x/c = -2.14$. A zero gradient condition for velocity was used at the outlet. The top wall was specified as a slip wall. The inlet $Re_\theta = 10,975$ was much larger than the training data. The Reynolds number based on channel height was 20,000.

Figure 4.17 shows that the modified model improved both the FPG and APG regions. The skin friction was under predicted over the bump and in the recovery region by the base model. The ML algorithm was able to differentiate and make the necessary corrections. Since the C_f was under predicted by the base model, a decrease in production of ω is needed to correct the eddy viscosity. Indeed, a correction factor of less than one was produced in these regions, although figure 4.17b



(a) C_f (b) Zones

Figure 4.17 Skin Friction Coefficient and zones for NASA Hump.

shows the largest correction to be on the lee side. The shortened reattachment length contributes to the increased C_f at the top of the bump.

4.4 Conclusion

The FIML methodology can be regarded as a means to extract variables that are peculiar to a turbulence closure model. In the present case, it provides a correction field $\beta(x, y)$ for the $k - \omega$ model. The data are patchy, with isolated zones of $\beta(x, y) > 1$ or $\beta(x, y) < 1$, within a field of $\beta(x, y) = 1$. That is a motive for developing a parametric dependence via a classifier, rather than parameterizing the β field by a continuous correlation function. The classifier bins the patches into ranges of the features. Here, a zonal model was created from data using a bag of decision trees as the classifier.

Since this setup is still exploratory, computational efficiency was not a priority. Computing the features, themselves, did not create any perceivable computational overhead. However, the bag of trees was not implemented in OpenFOAM, and an external function needed to be called, one cell at a time, which did increase the over all computation time. This, however, would not reflect the actual efficiency of the zonal method, once it had been optimized.

One important difference between the present, and previous studies, is that the features used to classify the zones are local (except for wall distance, which is used in most popular turbulence models), non dimensional, rotationally and translationally invariant (table 4.5). The dominant feature was deemed to be Q (figure 4.6), unlike previous studies, where it was wall-distance. The pressure gradient parameters also had a strong influence on the present classifier.

The zonal model shows as good, or better results than the base model in the test cases, tending to revert to the base model where that provides accurate predictions. All of this is done automatically, based on local features, and the decision tree. In that sense, the approach is capable of identifying and correcting regions of flow where the original model is likely to fail.

At this point, it is evident that zonal models can be developed with machine-learning methods, using data that are extracted by the FIML optimization. Further testing and training of the trees should be explored to make this approach more robust.

CHAPTER 5. SUMMARY

This work is an attempt to improve current RANS models using Machine learning. This is in line with the current trend to improve RANS models. To this end, a database was created using wall resolved LES. The database consisted of a series of bumps ranging from a geometry that had attached flow over APG to bumps that had separation of varying degrees. Following the data generation, an optimization algorithm was solved to determine optimal coefficients for the $k - \omega$ model to get a velocity field comparable to the LES velocity field. It was observed that the coefficients were changed near the wall showing where the RANS model needed correction. This correction was then generalized to be applicable to different flow geometries using ML which created a functional relation between a set of local features and the correction factor. It was shown that this approach was successful in improving RANS model predictions over several 2D flows.

APPENDIX A. ANN FOR TURBULENCE MODELING

An artificial neural net (ANN) can be used to form a functional relationship between field variables and the correction factor (β) as described in chapter 4. However, instead of having a classification function, a regression function was formulated using ANN giving rise to a continuous correction field. Thus, given some values of the features at each grid point, a correction factor would be generated.

As in the case with the bag of trees approach, pre-processing the data set was required. Data above the inlet momentum thickness was removed. This was done to prevent overwhelming the network with data that did not need correction.

To ensure the applicability of the model to a wide number of flows, the features were ensured to be coordinate and translation invariant (other than the wall distance). Features (table A.1) were again selected based on looking at the contours that showed high visual correlation with the β field.

The ANN corrections needed to be limited to over the bump and near the wall. This was done using the Spalart Shur tensor which can be used to detect streamline curvature. In 2D only one component of the tensor was used to determine if the ANN was to be switched "on" or "off", this was done to ensure that the ANN would only be activated over the bump. When the neural network was "off" the correction factor was set to 1. To ensure the smoothness of the β field, a "tanh"

Table A.1 Input Feature set for ANN.

#	*description*	$feature$	$non-dimensionalized\ feature$
1	magnitude of velocity gradient	$\|S\|$	$\frac{\|S\|L^2}{\nu}$
2	wall distance	d	$\frac{\sqrt{k}d}{\nu}$
3	turbulent viscosity	ν_t	$\frac{\nu_t}{\nu}$

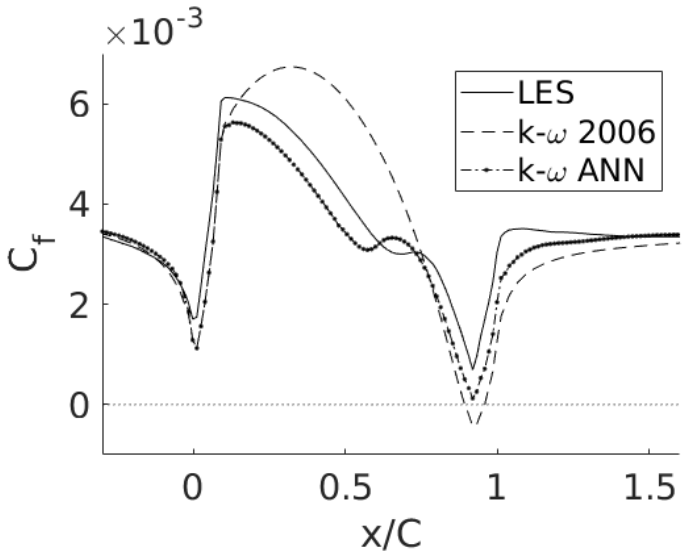
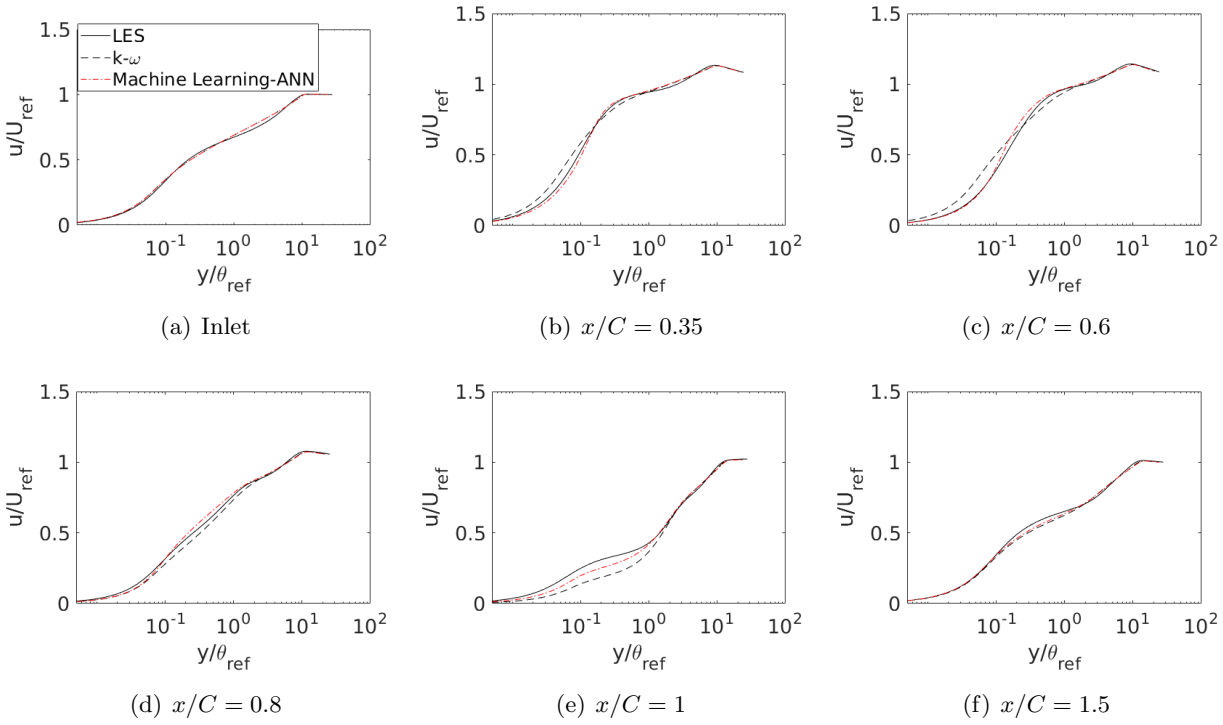


Figure A.1 Skin friction coefficient for h20.

blending function was used. In addition to the detection of the bump, two different ANNs were used, one for the FPG region and one for the APG region. A single ANN could be used but it was found that this created difficulty in being able to reproduce an ANN that could produce similar results. Again, a "tanh" function was used to ensure the smoothness of β in the transition between FPG and APG region. A single neural network with 8 neurons in the hidden layers and a sigmoid activation function was used. The mean squared error with the sum of weights (regularization) was used as the objective function for the ANN. All the computations were in OpenFOAM and the neural network matrices were implemented in the RANS model.

h20, described earlier, was used as the training case. Using the ANN on the same case yielded significant improvements. The original model predicted a spurious separation bubble which was no longer the case in the ANN augmented model. The FPG region prediction was also improved by decreasing the predicted skin friction and the prediction for the skin friction was improved overall. This is shown in the following figure A.1. The velocity profiles at various streamwise locations shown in figure A.2 demonstrate the improvement in overall velocity predictions by the neural network augmented turbulence model.



(a) Inlet (b) $x/C = 0.35$ (c) $x/C = 0.6$

(d) $x/C = 0.8$ (e) $x/C = 1$ (f) $x/C = 1.5$

Figure A.2 x-velocity at various x locations for h20.

A.1 Test Cases

A.1.1 h26

h26 was the second bump in the series of bumps. The flow at inlet $Re_\theta = 2500$ just started to separate with the C_f barely going below zero. However, the $k - \omega$ 2006 model predicted a large separation bubble. Using the ANN model the separation bubble was reduced significantly but the bubble was still bigger than LES predictions. The skin friction overall was improved as shown in figure A.3. The velocity profiles were also improved significantly as shown in figure A.4. This is most apparent at $x/C = 1$ where the spurious prediction of flow reversal was corrected. This shows that the neural network is capable of reversing the direction of the flow if needed.

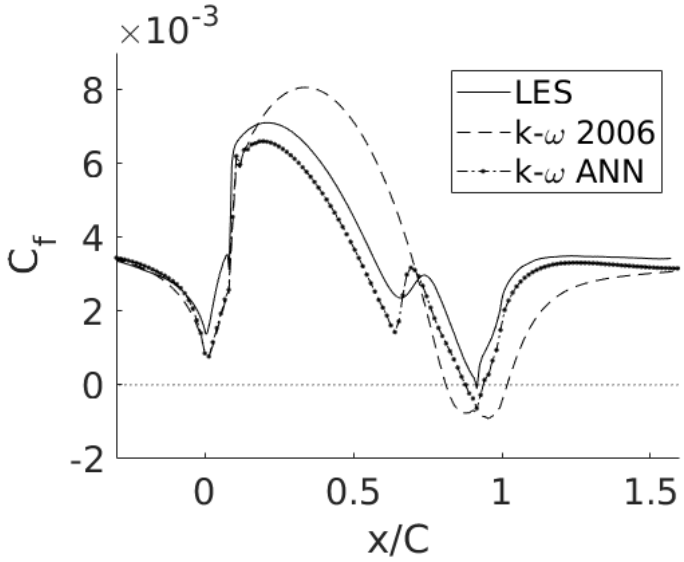
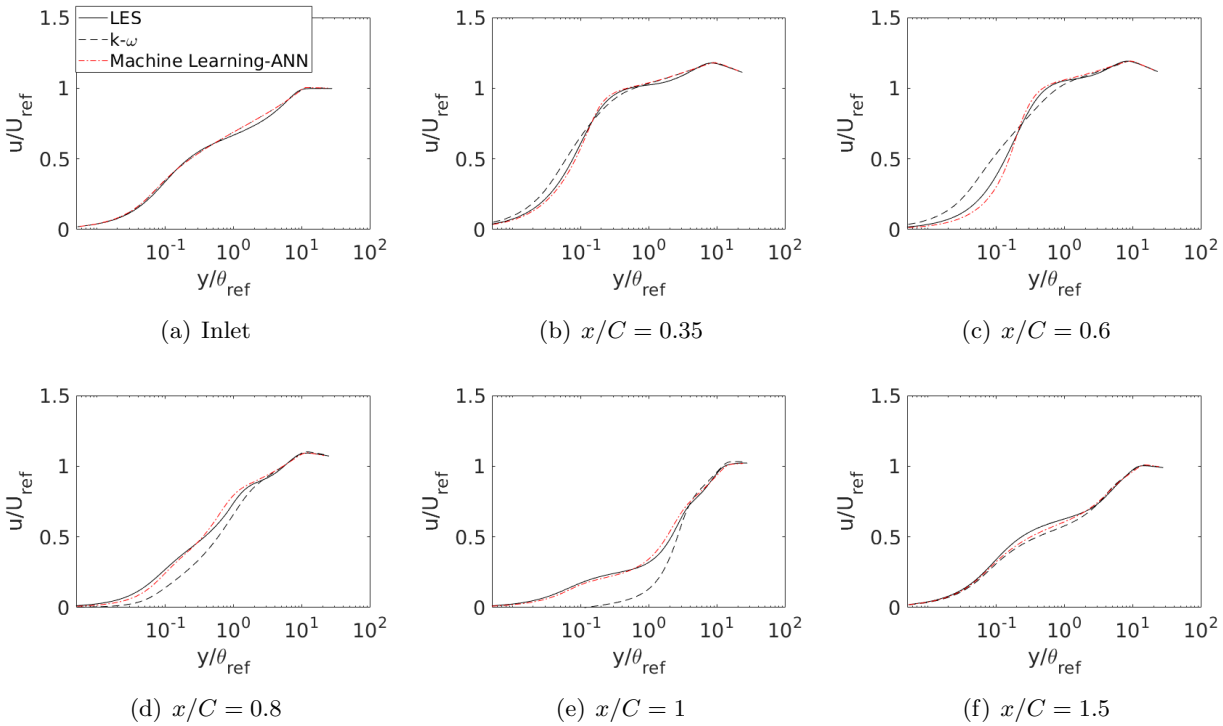


Figure A.3 Skin friction coefficient for h26.



(a) Inlet (b) $x/C = 0.35$ (c) $x/C = 0.6$

(d) $x/C = 0.8$ (e) $x/C = 1$ (f) $x/C = 1.5$

Figure A.4 x-velocity at various x locations for h26.

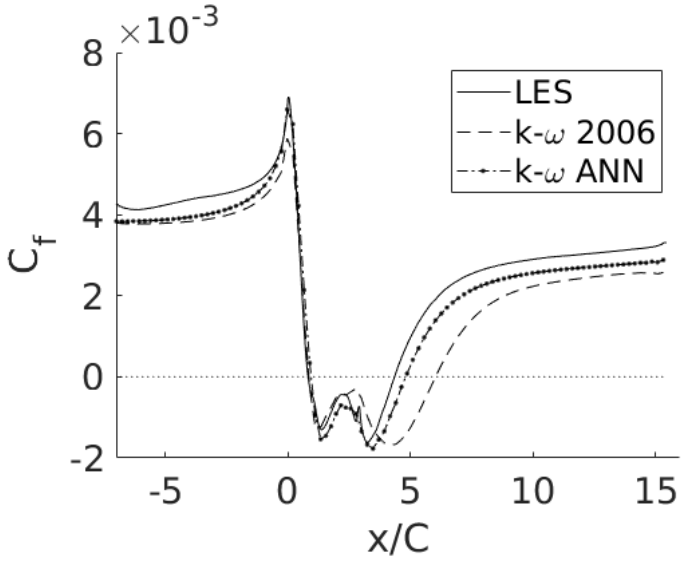


Figure A.5 Skin friction coefficient for curved back step.

A.1.2 Curved back step

The curved back step is another test case for turbulence models and combines the effect of APG and surface curvature. The base $k-\omega$ 2006 model over predicted the length of the separation bubble and the ANN model reduced the over prediction (figure A.5). The test case had an inlet $Re_\theta = 1190$ which was different than the training case demonstrating that the ANN was applicable to flows with a different Reynolds number than the training case. Note that this observation many not be universally applicable.

A.1.3 Limitations

Although these results are promising the various "tricks" needed to make the model work made it into a "not so easy to implement" model. The value of the Spalart Shur tensor component may not be universally applicable. The use of two neural networks and the requirement of a blending function made the model unnecessarily complicated. A single neural network was trained and found to be equally effective but such a network was found to be difficult to reproduce. That being said, this approach is definitly worth exploring further with more complex networks and other advanced ML techniques.

APPENDIX B. CONTINUOUS ADJOINT FORMULATION

The field inversion in chapter 3 and chapter 4 was performed using a optimization procedure which used the descrete adjoint formulation to calculate the gradient. This optimizer was kindly provided by Dr. Duraisamy. The optimizer could be modified to optimized coefficients in various models such as the $k-\omega$ or Spalart Allamaras model amongst others and was implemented in a finite volume C++ code ("Joe").

Another optimizer which used the continuous adjoint formulation to calculate the gradient was implemented in OpenFOAM. However instead of optimizing specific turbulence models, the eddy viscosity itself was optimized. The correction coefficient (β) was multiplied to the eddy viscosity. Following is the continuous adjoint formulation for obtaining the gradient for the eddy viscosity correction. The formulation closely follows Othmer (2008) where the reader can find a detailed derivation.

The problem statement could be stated as follows,

$$minimize, \quad J = \frac{\sum_\Omega (U - U^{LES})^2 V/d^2}{\sum_\Omega V/d^2} \tag{B.1}$$

where Ω stands for the flow domain, V for the cell volume and the cost function has been scaled by inverse wall distance (d) squared to emphasize on the near wall velocity. J is subjected to the constraints (steady state RANS equations):

$$\begin{aligned} (R_1, R_2, R_3)^T &= (\mathbf{v}\cdot\nabla)\mathbf{v} + \nabla p - \nabla\cdot(2(\nu+\beta\nu_T)D(\mathbf{v})) \\ R_4 &= -\nabla\cdot\mathbf{v} \end{aligned} \tag{B.2}$$

where $D(\mathbf{v}) = 1/2(\nabla\mathbf{v} + \nabla\mathbf{v}^T)$ is the strain rate tensor and $\mathbf{v}$ represents the velocity field and p the pressure field.

The constrained optimization problem can be solved by modifying the objective function by introducing Lagrange multipliers (adjoint variables, $\mathbf{v}_a$ and p_a).

$$L = J + \int_{\Omega} (\mathbf{v}_a, p_a)\mathbf{R}d\Omega \tag{B.3}$$

The sensitivity of the modified cost function w.r.t. β_i at each cell can then be written as,

$$\frac{\partial L}{\partial \beta_i} = \mathbf{v}_a \cdot (\nabla \cdot (2\nu_T \mathbf{D}(\mathbf{v}))) \cdot V_i \tag{B.4}$$

where V_i is the cell volume.

The adjoint equations can be derived and we get the following,

$$\begin{aligned} -\nabla \mathbf{v}_a \cdot \mathbf{v} - (\mathbf{v} \cdot \nabla)\mathbf{v}_a &= -\nabla p_a + \nabla \cdot (2(\nu + \beta\nu_T)D(\mathbf{v}_a)) - \tfrac{\partial J_\Omega}{\partial \mathbf{v}} \\ \nabla \cdot \mathbf{v}_a &= 0 \end{aligned} \tag{B.5}$$

The boundary conditions for the adjoint equations were as follows,

- the wall and inlet: $\mathbf{v}_a = 0$; $(n \cdot \nabla)p_a = 0$
- top: $(n \cdot \nabla)\mathbf{v}_a = 0$; $(n \cdot \nabla)p_a = 0$
- outlet: $\mathbf{v}_{at} = \frac{(\nu+\beta\nu_T)/\Delta \mathbf{v}_{at,neighbour}}{v_n + (\nu+\beta\nu_T)/\Delta}$; $p_a = \mathbf{v}_a \mathbf{v} + v_n v_{an} + (\nu + \beta\nu_T)(n \cdot \nabla)v_{an}$

Using the above adjoint equations to compute sensitivity of the cost function to β, the correction field (β) was updated and once the minimum had been found a RANS simulation using the final β yielded much improved skin friction and velocity profiles. Figure B.1 shows the improved C_f after the optimization, note that this is not as accurate as the results given by the discrete adjoint optimizer and there is room for improvement. Figure B.2 shows an improvement in velocity profiles over the original RANS model.

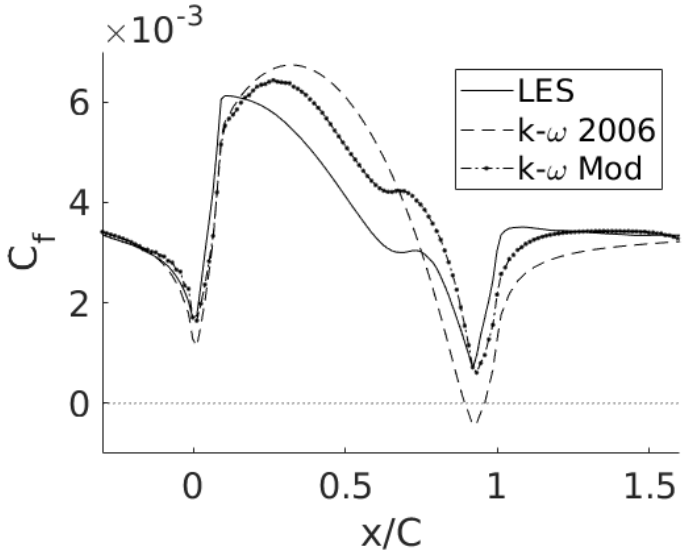
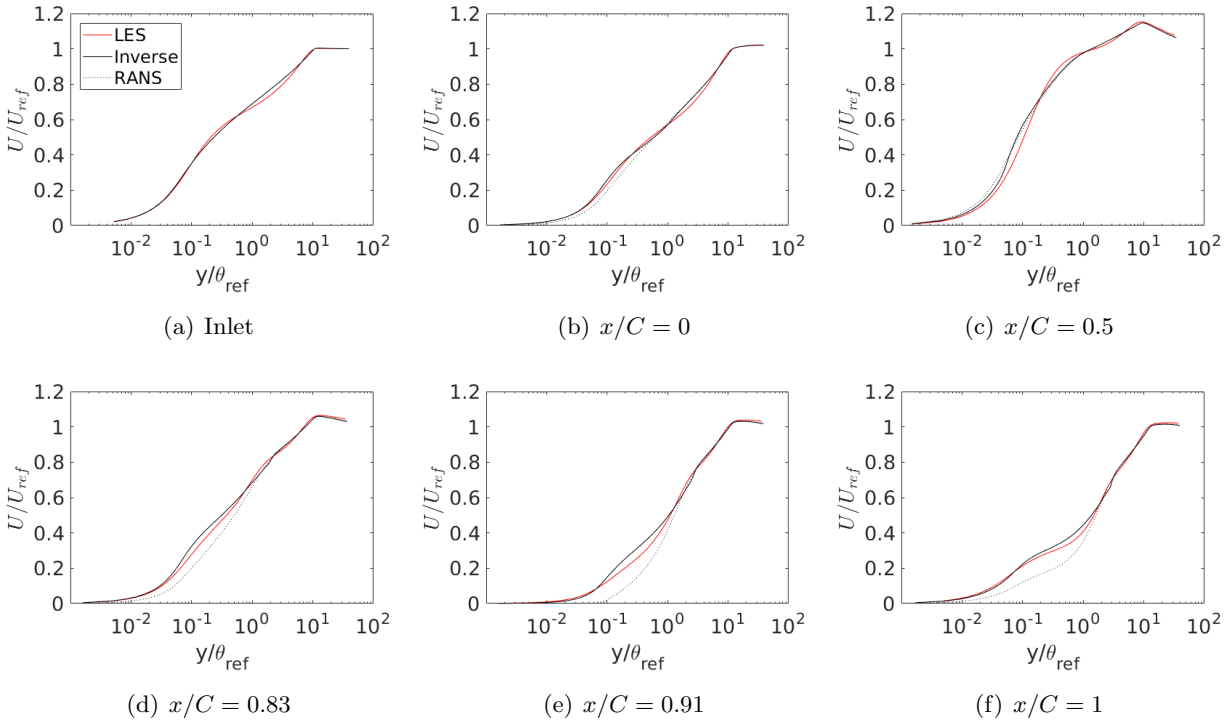


Figure B.1 Skin friction coefficient for h20.



(a) Inlet (b) $x/C = 0$ (c) $x/C = 0.5$

(d) $x/C = 0.83$ (e) $x/C = 0.91$ (f) $x/C = 1$

Figure B.2 x-velocity at various x locations for h20.

Bibliography

Alving, A. E., Smits, A. J., and Watmuff, J. H. (1990). Turbulent boundary layer relaxation from convex curvature. *Journal of Fluid Mechanics*, 211:529–556.

Arolla, S. K. (2016). Inflow turbulence generation for eddy-resolving simulations of turbomachinery flows. *Journal of Fluids Engineering*, 138(3):031201.

Bandyopadhyay, P. R. and Ahmed, A. (1993). Turbulent boundary layers subjected to multiple curvatures and pressure gradients. *Journal of Fluid Mechanics*, 246:503–527.

Baskaran, V., Smits, A., and Joubert, P. (1987). A turbulent flow over a curved hill part 1. growth of an internal boundary layer. *Journal of Fluid Mechanics*, 182:47–83.

Batchelor, G. K. (1967). *An Introduction to Fluid Dynamics*. Cambridge University press.

Blackwelder, R. F. and Kovasznay, L. S. (1972). Large-scale motion of a turbulent boundary layer during relaminarization. *Journal of Fluid Mechanics*, 53(1):61–83.

Breuer, M., Peller, N., Rapp, C., and Manhart, M. (2009). Flow over periodic hills - numerical and experimental study in a wide range of Reynolds numbers. *Computers and Fluids*, 38:433.

Davidson, P. (2015). *Turbulence: an introduction for scientists and engineers*. Oxford University Press.

Driver, D. M. and Seegmiller, H. L. (1985). Features of a reattaching turbulent shear layer in divergent channelflow. *AIAA journal*, 23(2):163–171.

Duraisamy, K., Zhang, Z. J., and Singh, A. P. (2015). New approaches in turbulence and transition modeling using data-driven techniques. In *53rd AIAA Aerospace Sciences Meeting*, page 1284.

Durbin, P. A. (2018). Some recent developments in turbulence closure modeling. *Ann. Rev. Fluid Mechanics*, 50:1–27.

Durbin, P. A. and Reif, B. P. (2011). *Statistical theory and modeling for turbulent flows*. John Wiley & Sons.

Gillis, J. and Johnston, J. (1983). Turbulent boundary-layer flow and structure on a convex wall and its redevelopment on a flat wall. *Journal of Fluid Mechanics*, 135:123–153.

Greenblatt, D. (2018). NASA Turbulence Modeling Resource. https://turbmodels.larc.nasa.gov/nasahump_grids.html.

Jesus, A., Azevedo, J. L., and Laval, J.-P. (2013). Large eddy simulations and RANS computations of adverse pressure gradient flows. In *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, page 267.

Kitsios, V., Atkinson, C., Sillero, J. A., Borrell, G., Gungor, A. G., Jiménez, J., and Soria, J. (2016). Direct numerical simulation of a self-similar adverse pressure gradient turbulent boundary layer. *International Journal of Heat and Fluid Flow*, 61:129–136.

Kuban, L., Laval, J.-P., Elsner, W., Tyliszczak, A., and Marquillie, M. (2012). Les modeling of converging-diverging turbulent channel flow. *Journal of Turbulence*, (13):N11.

Lilly, D. K. (1992). A proposed modification of the germano subgrid-scale closure method. *Physics of Fluids A: Fluid Dynamics*, 4(3):633–635.

Ling, J., Jones, R., and Templeton, J. (2016a). Machine learning strategies for systems with invariance properties. *Journal of Computational Physics*, 318:22–35.

Ling, J., Kurzawski, A., and Templeton, J. (2016b). Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics*, 807:155–166.

Ling, J. and Templeton, J. (2015). Evaluation of machine learning algorithms for prediction of regions of high reynolds averaged navier stokes uncertainty. *Physics of Fluids*, 27(8):085103.

Matai, R. and Durbin, P. A. (2018). LES of turbulent flow over a parametric set of bumps. *To be s.* submitted.

Mollicone, J.-P., Battista, F., Gualtieri, P., and Casciola, C. (2017). Effect of geometry and reynolds number on the turbulent separated flow behind a bulge in a channel. *Journal of Fluid Mechanics*, 823:100–133.

Naughton, J. W., Viken, S., and Greenblatt, D. (2006). Skin friction measurements on the nasa hump model. *AIAA journal*, 44(6):1255–1265.

Obi, S., Aoki, K., and Masuda, S. (1993). Experimental and computational study of turbulent separating flow in an asymmetric plane diffuser. In *Ninth Symposium on Turbulent Shear Flows, Hyoto, Japan*, pages 305–1.

Othmer, C. (2008). A continuous adjoint formulation for the computation of topological and surface sensitivities of ducted flows. *International Journal for Numerical Methods in Fluids*, 58(8):861–877.

Page, F., Schlinger, W., Breaux, D., and Sage, B. (1952). Temperature gradients in turbulent gas streams. point values of eddy conductivity and viscosity. in uniform flow between parallel plates. *Industrial & Engineering Chemistry*, 44(2):424–430.

Parish, E. J. and Duraisamy, K. (2016a). A paradigm for data-driven predictive modeling using field inversion and machine learning. *Journal of Computational Physics*, 305:758 – 774.

Parish, E. J. and Duraisamy, K. (2016b). A paradigm for data-driven predictive modeling using field inversion and machine learning. *Journal of Computational Physics*, 305:758–774.

Parneix, S., Durbin, P., and Behnia, M. (1998). Computation of 3-d turbulent boundary layers using the v2f model. *Flow, Turbulence and Combustion*, 60(1):19–46.

Patel, V. (1965). Calibration of the preston tube and limitations on its use in pressure gradients. *Journal of Fluid Mechanics*, 23(1):185–208.

Patel, V. C. and Head, M. R. (1968). Reversion of turbulent to laminar flow. *Journal of Fluid Mechanics*, 34(2):371392.

Piomelli, U. (2014). Large eddy simulations in 2030 and beyond. *Phil. Trans. R. Soc. A*, 372(2022):20130320.

Pope, S. B. (2001). Turbulent flows.

Rumsey, C. (2018). NASA Turbulence Modeling Resource. https://turbmodels.larc.nasa.gov/Other_LES_Data/family_of_bumps.html.

Schiavo, L. A. A., Jesus, A. B., Azevedo, J. L. F., and Wolf, W. R. (2015). Adverse pressure gradient effects in the turbulent kinetic energy budget for channel flows. In *53rd AIAA Aerospace Sciences Meeting*, page 1276.

Seifert, A. and Pack, L. G. (2002). Active flow separation control on wall-mounted hump at high reynolds numbers. *AIAA Journal*, 40(7):1363–1372.

Singh, A. P., Matai, R., Mishra, A., Duraisamy, K., and Durbin, P. A. (2017). Data-driven augmentation of turbulence models for adverse pressure gradient flows. In *23rd AIAA Computational Fluid Dynamics Conference*, pages AIAA–2017–3626.

Slotnick, J., Khodadoust, A., Alonso, J., Darmofal, D., Gropp, W., Lurie, E., and Mavriplis, D. (2014). CFD vision 2030 study: A path to revolutionary computational aerosciences. Technical Report CR-2014-218178, NASA.

Spalart, P. R. (1988). Direct simulation of a turbulent boundary layer up to rθ= 1410. *Journal of fluid mechanics*, 187:61–98.

Spalart, P. R. and Watmuff, J. H. (1993). Experimental and numerical study of a turbulent boundary layer with pressure gradients. *Journal of Fluid Mechanics*, 249:337–371.

Tennekes, H., Lumley, J. L., Lumley, J., et al. (1972). *A first course in turbulence*. MIT press.

Tracey, B. D., Duraisamy, K., and Alonso, J. J. (2015). A machine learning strategy to assist turbulence model development. In *53rd AIAA Aerospace Sciences Meeting*, page 1287.

Tsuji, Y. and Morikawa, Y. (1976). Turbulent boundary layer with pressure gradient alternating in sign. *Aeronautical Quarterly*, 27(1):1528.

Wang, J.-X., Wu, J.-L., and Xiao, H. (2017). Physics-informed machine learning approach for reconstructing reynolds stress modeling discrepancies based on dns data. *Physical Review Fluids*, 2(3):034603.

Weatheritt, J. and Sandberg, R. (2016). A novel evolutionary algorithm applied to algebraic modifications of the rans stress–strain relationship. *Journal of Computational Physics*, 325:22–37.

Weatheritt, J. and Sandberg, R. (2017). The development of algebraic stress models using a novel evolutionary algorithm. *International Journal of Heat and Fluid Flow*, 68:298–318.

Webster, D., DeGraaff, D., and Eaton, J. (1996). Turbulence characteristics of a boundary layer over a two-dimensional bump. *Journal of Fluid Mechanics*, 320:53–69.

Weller, H. G., Tabor, G., Jasak, H., and Fureby, C. (1998). A tensorial approach to computational continuum mechanics using object-oriented techniques. *Computers in physics*, 12(6):620–631.

Wilcox, D. C. (2008). Formulation of the k–x turbulence model revisited. *AIAA J*, 46(11):2823–2838.

Wu, J.-L., Wang, J.-X., Xiao, H., and Ling, J. (2017). A priori assessment of prediction confidence for data-driven turbulence modeling. *Flow, Turbulence and Combustion*, 99(1):25–46.

Wu, X. and Squires, K. (1998a). Prediction of the three-dimensional turbulent boundary layer over a swept bump. *AIAA Journal*, 36(4):505–514.

Wu, X. and Squires, K. D. (1998b). Numerical investigation of the turbulent boundary layer over a bump. *Journal of Fluid Mechanics*, 362:229–271.

Yarlanki, S., Rajendran, B., and Hamann, H. (2012). Estimation of turbulence closure coefficients for data centers using machine learning algorithms. In *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2012 13th IEEE Intersociety Conference on*, pages 38–42. IEEE.